

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
3 October 2002 (03.10.2002)

PCT

(10) International Publication Number
WO 02/077267 A2

(51) International Patent Classification⁷: **C12Q**

(21) International Application Number: PCT/US02/09611

(22) International Filing Date: 27 March 2002 (27.03.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/279,228 27 March 2001 (27.03.2001) US
60/291,501 15 May 2001 (15.05.2001) US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

US 60/279,228 (CIP)
Filed on 27 March 2001 (27.03.2001)
US 60/291,501 (CIP)
Filed on 15 May 2001 (15.05.2001)

(71) Applicant (for all designated States except US): **CORVAS INTERNATIONAL, INC.** [US/US]; 3030 Science Park Road, San Diego, CA 92121 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **MADISON, Edwin, L.** [US/US]; 11005 Cedarcrest Way, San Diego, CA 92121 (US). **ONG, Edgar, O.** [CA/US]; 10738 Glendover Lane, San Diego, CA 92126 (US).

(74) Agents: **SEIDMAN, Stephanie, L.** et al.; Heller Ehrman White & McAuliffe LLP, 4350 La Jolla Village Drive, San Diego, CA 92122-1246 (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),

Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

— as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii)) for the following designations AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZM, ZW, ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)

— as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii)) for the following designations AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZM, ZW, ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)

Published:

— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: NUCLEIC ACID MOLECULES ENCODING A TRANSMEMBRAN SERINE PROTEASE 9, THE ENCODED POLYPEPTIDES AND METHODS BASED THEREON

(57) Abstract: Provided herein are type II transmembrane serine protease 9 (MTSP9) polypeptides. Zymogen and activated forms of these polypeptides as well as single and two chain forms of the protease domain are also provided. Methods using the polypeptides to identify compounds that modulate the protease activity of an MTSP9 are provided.



WO 02/077267 A2

-1-

**NUCLEIC ACID MOLECULES ENCODING A TRANSMEMBRANE SERINE
PROTEASE 9, THE ENCODED POLYPEPTIDES AND METHODS BASED THEREON
RELATED APPLICATIONS**

Benefit of priority is claimed to U.S. provisional application Serial No.
5 60/279,228, filed March 27, 2001, Edwin L. Madison and Edgar O. Ong,
entitled "NUCLEIC ACID MOLECULES ENCODING TRANSMEMBRANE SERINE
PROTEASE 9, THE ENCODED PROTEINS AND METHODS BASED THEREON",
and to U.S. provisional application Serial No. 60/291,501, filed May 15, 2001,
to Edwin L. Madison and Edgar O. Ong, entitled "NUCLEIC ACID MOLECULES
10 ENCODING TRANSMEMBRANE SERINE PROTEASE 9, THE ENCODED PROTEINS
AND METHODS BASED THEREON." Where permitted, the subject matter of
each of these application is incorporated in its entirety by reference thereto.

FIELD OF INVENTION

Nucleic acid molecules that encode proteases and portions thereof,
15 particularly protease domains are provided. Also provided are prognostic,
diagnostic and therapeutic methods using the proteases and domains thereof and
the encoding nucleic acid molecules.

BACKGROUND OF THE INVENTION AND OBJECTS THEREOF

Cancer, which is a leading cause of death in the United States, is
20 characterized by an increase in the number of abnormal neoplastic cells, which
proliferate to form a tumor mass, the invasion of adjacent tissues by these
neoplastic tumor cells, and the generation of malignant cells that metastasize via
the blood or lymphatic system to regional lymph nodes and to distant sites.
Among the hallmarks of cancer is a breakdown in the communication among
25 tumor cells and their environment. Normal cells do not divide in the absence of
stimulatory signals and cease dividing in the presence of inhibitory signals.
Growth-stimulatory and growth-inhibitory signals, are routinely exchanged
between cells within a tissue. In a cancerous, or neoplastic, state, a cell
acquires the ability to "override" these signals and to proliferate under conditions
30 in which normal cells do not grow.

-2-

In order to proliferate tumor cells acquire a number of distinct aberrant traits reflecting genetic alterations. The genomes of certain well-studied tumors carry several different independently altered genes, including activated oncogenes and inactivated tumor suppressor genes. Each of these genetic changes appears to be responsible for imparting some of the traits that, in the aggregate, represent the full neoplastic phenotype.

A variety of biochemical factors have been associated with different phases of metastasis. Cell surface receptors for collagen, glycoproteins such as laminin, and proteoglycans, facilitate tumor cell attachment, an important step in invasion and metastases. Attachment triggers the release of degradative enzymes which facilitate the penetration of tumor cells through tissue barriers. Once the tumor cells have entered the target tissue, specific growth factors are required for further proliferation. Tumor invasion and progression involve a complex series of events, in which tumor cells detach from the primary tumor, break down the normal tissue surrounding it, and migrate into a blood or lymphatic vessel to be carried to a distant site. The breaking down of normal tissue barriers is accomplished by the elaboration of specific enzymes that degrade the proteins of the extracellular matrix that make up basement membranes and stromal components of tissues.

A class of extracellular matrix degrading enzymes has been implicated in tumor invasion. Among these are the matrix metalloproteinases (MMP). For example, the production of the matrix metalloproteinase stromelysin is associated with malignant tumors with metastatic potential (see, *e.g.*, McDonnell *et al.* (1990) *Smnrs. in Cancer Biology* 1:107-115; McDonnell *et al.* (1990) *Cancer and Metastasis Reviews* 9:309-319).

The capacity of cancer cells to metastasize and invade tissue is facilitated by degradation of the basement membrane. Several proteinase enzymes, including the MMPs, have been reported to facilitate the process of invasion of tumor cells. MMPs are reported to enhance degradation of the basement membrane, which thereby permits tumorous cells to invade tissues. For example, two major metalloproteinases having molecular weights of about 70 kDa and 92 kDa appear to enhance ability of tumor cells to metastasize.

Type II Transmembrane Serine Proteases (TTSPs)

In addition to the MMPs, serine proteases have been implicated in neoplastic disease progression. Most serine proteases, which are either secreted enzymes or are sequestered in cytoplasmic storage organelles, have roles in blood coagulation, wound healing, digestion, immune responses and tumor invasion and metastasis. A class of cell surface proteins designated type II transmembrane serine proteases, which are membrane-anchored proteins with additional extracellular domains, has been identified. As cell surface proteins, they are positioned to play a role in intracellular signal transduction and in mediating cell surface proteolytic events.

Cell surface proteolysis is a mechanism for the generation of biologically active proteins that mediate a variety of cellular functions. Membrane-associated proteases include membrane-type metalloproteinases (MT-MMP), ADAMs (proteases that contain disintegrin-like and metalloproteinase domains) and the TTSPs. In mammals, at least 17 members of the TTSP family are known, including seven in humans (see, Hooper *et al.* (2001) *J. Biol. Chem.* 276:857-860). These include: corin (accession nos. AF133845 and AB013874; see, Yan *et al.* (1999) *J. Biol. Chem.* 274:14926-14938; Tomia *et al.* (1998) *J. Biochem.* 124:784-789; Uan *et al.* (2000) *Proc. Natl. Acad. Sci. U.S.A.* 97:8525-8529); enterpeptidase (also designated enterokinase; accession no. U09860 for the human protein; see, Kitamoto *et al.* (1995) *Biochem.* 27: 4562-4568; Yahagi *et al.* (1996) *Biochem. Biophys. Res. Commun.* 219:806-812; Kitamoto *et al.* (1994) *Proc. Natl. Acad. Sci. U.S.A.* 91:7588-7592; Matsushima *et al.* (1994) *J. Biol. Chem.* 269:19976-19982); human airway trypsin-like protease (HAT; accession no. AB002134; see Yamaoka *et al.* *J. Biol. Chem.* 273:11894-11901); MTSP1 and matriptase (also called TADG-15; see SEQ ID Nos. 1 and 2; accession nos. AF133086/AF118224, AF04280022; Takeuchi *et al.* (1999) *Proc. Natl. Acad. Sci. U.S.A.* 96:11054-11161; Lin *et al.* (1999) *J. Biol. Chem.* 274:18231-18236; Takeuchi *et al.* (2000) *J. Biol. Chem.* 275:26333-26342; and Kim *et al.* (1999) *Immunogenetics* 49:420-429); hepsin (see, accession nos. M18930, AF030065, X70900; Leytus *et al.* (1988) *Biochem.* 27: 11895-11901; Vu *et al.* (1997) *J. Biol. Chem.* 272:31315-31320;

-4-

and Farley *et al.* (1993) *Biochem. Biophys. Acta* 1173:350-352; and see, U.S. Patent No. 5,972,616); TMPRS2 (see, Accession Nos. U75329 and AF113596; Paoloni-Giacobino *et al.* (1997) *Genomics* 44:309-320; and Jacquinet *et al.* (2000) *FEBS Lett.* 468: 93-100); and TMPRSS4 (see, Accession No. NM
5 016425; Wallrapp *et al.* (2000) *Cancer* 60:2602-2606).

Serine proteases, including transmembrane serine proteases and secreted proteases, have been implicated in processes involved in neoplastic development and progression. While the precise, detailed mechanism by which these proteases promote tumor growth and progression has not been elaborated,
10 serine proteases and inhibitors thereof are involved in the control of many intra- and extracellular physiological processes, including degradative actions in cancer cell invasion, metastatic spread, and neovascularization of tumors, that are involved in tumor progression. It is believed that proteases are involved in the degradation of extracellular matrix (ECM) and contribute to tissue remodeling,
15 and are necessary for cancer invasion and metastasis. The activity and/or expression of some proteases have been shown to correlate with tumor progression and development.

For example, a membrane-type serine protease MTSP1 (also called matriptase; see SEQ ID Nos. 1 and 2 from U.S. Patent No. 5,972,616; and
20 GenBank Accession No. AF118224; (1999) *J. Biol. Chem.* 274:18231-18236; U.S. Patent No. 5,792,616; see, also Takeuchi (1999) *Proc. Natl. Acad. Sci. U.S.A.* 96:11054-1161) that is expressed in epithelial cancer and normal tissue (Takeuchi *et al.* (1999) *Proc. Natl. Acad. Sci. USA* 96:11054-61) has been identified. Matriptase was originally identified in human breast cancer cells as a
25 major gelatinase (see, U.S. Patent No. 5,482,848) and was initially believed to be a type of matrix metalloprotease (MMP). It has been proposed that it plays a role in the metastasis of breast cancer. Matriptase also is expressed in a variety of epithelial tissues with high levels of activity and/or expression in the human gastrointestinal tract and the prostate. MTSPs, designated MTSP3, MTSP4,
30 MTSP6 have been described in published International PCT application No. WO 01/57194, based in International PCT application No. PCT/US01/03471.

-5-

Prostate-specific antigen (PSA), a kallikrein-like serine protease, degrades extracellular matrix glycoproteins fibronectin and laminin, and, has been postulated to facilitate invasion by prostate cancer cells (Webber *et al.* (1995) *Clin. Cancer Res.*, 1(10):1089-94). Blocking PSA proteolytic activity with

- 5 PSA-specific monoclonal antibodies results in a dose-dependent decrease *in vitro* in the invasion of the reconstituted basement membrane Matrigel by LNCaP human prostate carcinoma cells which secrete high levels of PSA.

Hepsin, a cell surface serine protease identified in hepatoma cells, is overexpressed in ovarian cancer (Tanimoto *et al.* (1997) *Cancer Res.*, 57(14):2884-7). The hepsin transcript appears to be abundant in carcinoma tissue and is almost never expressed in normal adult tissue, including normal ovary. It has been suggested that hepsin is frequently overexpressed in ovarian tumors and therefore can be a candidate protease in the invasive process and growth capacity of ovarian tumor cells.

- 15 A serine protease-like gene, designated normal epithelial cell-specific 1 (NES1) (Liu *et al.*, *Cancer Res.*, 56(14):3371-9 (1996)) has been identified. Although expression of the NES1 mRNA is observed in all normal and immortalized nontumorigenic epithelial cell lines, the majority of human breast cancer cell lines show a drastic reduction or a complete lack of its expression.
- 20 The structural similarity of NES1 to polypeptides known to regulate growth factor activity and a negative correlation of NES1 expression with breast oncogenesis suggest a direct or indirect role for this protease-like gene product in the suppression of tumorigenesis.

Hence transmembrane serine proteases appear to be involved in the etiology and pathogenesis of tumors. There is a need to further elucidate their role in these processes and to identify additional transmembrane proteases.

Therefore, it is an object herein to provide transmembrane serine protease (MTSP) proteins and nucleic acids encoding such MTSP proteases that are involved in the regulation of or participate in tumorigenesis and/or

- 30 carcinogenesis. It is also an object herein to provide prognostic, diagnostic and therapeutic screening methods using such proteases and the nucleic acids encoding such proteases.

SUMMARY OF THE INVENTION

Provided herein are members of the Transmembrane Serine Protease family, particularly the Type II Transmembrane Serine Protease (TTSP) family (also referred to herein as MTSPs), and more particularly TTSP family members
5 whose functional activity and/or expression differs in tumor cells from non-tumor cells in the same tissue. The MTSP provided herein is a MTSP family member, designated herein as MTSP9. The protease domain and full-length protein, including the zymogen and activated forms, and uses thereof are also provided. Proteins encoded by splice variants are also provided.

10 Assays for identifying effectors, such as compounds, including small molecules, and conditions, such pH, temperature and ionic strength, that modulate the activation, expression or activity of MTSP9 are also provided herein. In exemplary assays, the effects of test compounds on the ability of a protease domain of MTSP9 to proteolytically cleave a known substrate, typically
15 a fluorescently, chromogenically or otherwise detectably labeled substrate, are assessed. Agents, generally compounds, particularly small molecules, that modulate the activity of the protease domain are candidate compounds for modulating the activity of the MTSP9. The protease domains can also be used to produce protease-specific antibodies. The protease domains provided herein
20 include, but are not limited to, the single chain region having an N-terminus at the cleavage site for activation of the zymogen, through the C-terminus, or C-terminal truncated portions thereof that exhibit proteolytic activity as a single-chain polypeptide in *vitro* proteolysis assays, of any MTSP family member, including MTSP9, generally from a mammal, including human, that, for example,
25 is expressed in tumor cells at different levels from non-tumor cells.

Nucleic acid molecules encoding the proteins and protease domains are also provided. Nucleic acid molecules that encode a single-chain protease domain or catalytically active portion thereof and also those that encode the full-length MTSP9 are provided. Nucleic acid encoding the protease domain
30 (nucleotides 31-729) and upstream nucleic acid in SEQ ID No. 5; and the protease domain of MTSP9 is set forth in SEQ ID No. 6 (amino acids 11-232)

-7-

and in SEQ ID No. 16. The protein sequence and encoding nucleic acid sequence of the full-length MTSP9 are set forth in SEQ ID Nos. 18 and 17.

Also provided are nucleic acid molecules that hybridize to such MTSP9-encoding nucleic acid along their full-length or along at least about 70%, 80% or
5 90% of the full-length and encode the protease domain or portion thereof are provided. Hybridization is generally effected under conditions of at least low, generally at least moderate, and often high stringency.

The isolated nucleic acid fragment is DNA, including genomic or cDNA, or is RNA, or can include other components, such as protein nucleic acid or other
10 nucleotide analogs. The isolated nucleic acid may include additional components, such as heterologous or native promoters, and other transcriptional and translational regulatory sequences, these genes may be linked to other genes, such as reporter genes or other indicator genes or genes that encode indicators.

15 Also provided is an isolated nucleic acid molecule that includes the sequence of molecules that is complementary to the nucleotide sequence encoding MTSP9 or the portion thereof.

Also provided are fragments thereof or oligonucleotides that can be used as probes or primers and that contain at least about 10, 14, 16 nucleotides,
20 generally less than 1000 or less than or equal to 100, set forth in SEQ ID No. 5 or 17 (or the complement thereof); or contain at least about 30 nucleotides (or the complement thereof) or contain oligonucleotides that hybridize along their full-length (or at least about 70, 80 or 90% thereof) to any such fragments or oligonucleotides. The length of the fragments are a function of the purpose for
25 which they are used and/or the complexity of the genome of interest. Generally probes and primers contain less than about 50, 150 or 500 nucleotides.

Also provided are plasmids containing any of the nucleic acid molecules provided herein. Cells containing the plasmids are also provided. Such cells include, but are not limited to, bacterial cells, yeast cells, fungal cells, plant cells,
30 insect cells and animal cells.

Also provided is a method of producing MTSP9 by growing the above-described cells under conditions whereby the MTSP9 is expressed by the cells,

-8-

and recovering the expressed MTSP9 polypeptide. Methods for isolating nucleic acid encoding other MTSP9s are also provided.

Also provided are cells, generally eukaryotic cells, such as mammalian cells and yeast cells, in which the MTSP9 polypeptide is expressed on the surface of the cells. Such cells are used in drug screening assays to identify compounds that modulate the activity of the MTSP9 polypeptide. These assays, including *in vitro* binding assays, and transcription based assays in which signal transduction mediated directly or indirectly, such as via activation of pro-growth factors, by the MTSP9 is assessed.

Also provided are peptides that are encoded by such nucleic acid molecules. Included among those polypeptides are the MTSP9 protease domain or a polypeptide with amino acid changes such that the specificity and/or protease activity remains substantially unchanged. In particular, a substantially purified mammalian MTSP9 polypeptide is provided that includes a serine protease catalytic domain and may additionally include other domains. The MTSP9 can form homodimers and can also form heterodimers with some other protein, such as a membrane-bound protein. Also provided is a substantially purified protein including a sequence of amino acids that has at least 60%, 70%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% identity to the MTSP9 where the percentage identity is determined using standard algorithms and gap penalties that maximize the percentage identity. A human MTSP9 polypeptide is exemplified, although other mammalian MTSP9 polypeptides are contemplated. Splice variants of the MTSP9, particularly those with a proteolytically active protease domain, are contemplated herein.

In other embodiments, substantially purified polypeptides that include a protease domain of a MTSP9 polypeptide or a catalytically active portion thereof, but that do not include the entire sequence of amino acids set forth in SEQ ID No. 18 are provided. Among these are polypeptides that include a sequence of amino acids that has at least 60%, 70%, 80%, 85%, 90%, 95% or 100% sequence identity to SEQ ID No. 16 or 18.

In a specific embodiment, a nucleic acid that encodes a MTSP, designated MTSP9 is provided. In particular, the nucleic acid includes the sequence of nucleotides set forth in SEQ ID No. 5, particularly set forth as nucleotides 31-729 of SEQ ID No. 5, or SEQ ID No. 17 or a portion thereof that encodes a catalytically active polypeptide.

Also provided are nucleic acid molecules that hybridize under conditions of at least low stringency, generally moderate stringency, more typically high stringency to the SEQ ID No. 5 or 17 or degenerates thereof.

In one embodiment, the isolated nucleic acid fragment hybridizes to a nucleic acid molecule containing the nucleotide sequence set forth in SEQ ID No: 5 or 17 (or degenerates thereof) under high stringency conditions, in one embodiment contains the sequence of nucleotides set forth in SEQ ID Nos. 5 and 17. A full-length MTSP9 is set forth in SEQ ID No. 18 and is encoded by SEQ ID No. 17 or degenerates thereof.

Also provided are muteins of the single chain protease domain of MTSP9 particularly muteins in which the Cys residue in the protease domain that is free (*i.e.*, does not form disulfide linkages with any other Cys residue in the protease domain) is substituted with another amino acid substitution, typically, although not necessarily, with a conservative amino acid substitution or a substitution that does not eliminate the activity, and muteins in which a glycosylation site(s) is eliminated.

Hence muteins in which one or more of the Cys residues, particularly, a residue that is paired in the activated two form, but unpaired in the protease domain alone (*i.e.*, the Cys a residue position 26 (see SEQ ID Nos. 5, 6 and 16) in the protease domain), is/are replaced with any amino acid, typically, although not necessarily, a conservative amino acid residue, such as Ser, are contemplated. Muteins of MTSP9, particularly those in which Cys residues, such as the unpaired Cys in the single chain protease domain, is replaced with another amino acid that does not eliminate the activity, are provided. Muteins in which other

-10-

conservative or non-conservative amino acid substitutions in which catalytic activity is retained are also contemplated (see, *e.g.*, Table 1, for exemplary amino acid substitutions).

MTSP9 polypeptides, including, but not limited to splice variants thereof, and nucleic acids encoding MTSPs, and domains, derivatives and analogs thereof are provided herein. Single chain protease domains that have an N-terminus functionally equivalent to that generated by activation of the zymogen form of MTSP9 are also provided. The cleavage site for the protease domain of MTSP9 is between amino acid R₁₈₆ and amino acid I₁₈₇ (R↓IASG). There are two potential glycosylation sites at N₁₅₃ and N₃₀₃. Disulfide bonds form between the Cys residues C₁₇₅-C₂₉₂ to link the protease domain to another domain so that upon cleavage the resulting polypeptide is a two chain molecule. There are potential disulfide bonds as follows: C₂₁₂-C₂₂₈, C₃₃₇-C₃₅₃ and C₃₆₄-C₃₉₃. Hence C₂₉₂ is a free Cys in the single chain form of the protease domain, which can also be provided as a two chain molecule. It is shown herein, however, that the single chain and two chain forms are proteolytically active.

Hence, provided herein is a the family of transmembrane serine protease (MTSP) proteins designated MTSP9, and functional domains, especially protease (or catalytic) domains thereof, muteins and other derivatives and analogs thereof. Also provided herein are nucleic acids encoding the MTSP9s.

MTSPs are of interest because they appear to be expressed and/or activated at different levels in tumor cells from normal cells, or have functional activity that is different in tumor cells from normal cells, such as by an alteration in a substrate therefor, or a cofactor. MTSP9 is of interest because it is expressed or is active in tumor cells. Hence the MTSPs provided herein can serve as diagnostic markers for certain tumors.

Of interest herein are MTSPs that are expressed or are activated in certain tumor or cancer cells such lung, prostate, colon and breast cancers. In particular, it is shown herein, that MTSP9 is expressed and/or activated in a variety of tumor cells, including, for example, in lung carcinoma, leukemia and cervical carcinoma as well as in certain normal cells and tissues (see *e.g.*, EXAMPLES for tissue-specific expression profile). MTSP9 also can be a marker

-11-

for breast, prostate and colon cancer. The expression and/or activation of MTSP9 on or in the vicinity of a cell or in a bodily fluid in a subject can be a marker for breast, prostate, lung, colon and other cancers.

In certain embodiments, the MTSP9 polypeptide is detectable in a body fluid at a level that differs from its level in body fluids in a subject not having a tumor. In other embodiments, the polypeptide is present in a tumor; and a substrate or cofactor for the polypeptide is expressed at levels that differ from its level of expression in a non-tumor cell in the same type of tissue. In other embodiments, the level of expression and/or activity of the MTSP9 polypeptide in tumor cells differs from its level of expression and/or activity in non-tumor cells. In other embodiments, the MTSP9 is present in a tumor; and a substrate or cofactor for the MTSP9 is expressed at levels that differ from its level of expression in a non-tumor cell in the same type of tissue.

Also provided are methods for screening for compounds that modulate the activity of MTSP9. The compounds are identified by contacting them with the MTSP9 or protease domain thereof and a substrate for the MTSP9. A change in the amount of substrate cleaved in the presence of the compounds compared to that in the absence of the compound indicates that the compound modulates the activity of the MTSP9. Such compounds are selected for further analyses or for use to modulate the activity of the MTSP9, such as inhibitors or agonists. The compounds can also be identified by contacting the substrates with a cell that expresses the MTSP9 or the extracellular domain or proteolytically active portion thereof.

Also provided herein are methods of modulating the activity of the MTSP9 and screening for compounds that modulate, including inhibit, antagonize, agonize or otherwise alter the activity of the MTSP9. Of particular interest is the extracellular domain of MTSP9 that includes the proteolytic (catalytic) portion of the protein.

Cells, combinations, kits and articles of manufacture containing the MTSP9 polypeptides, domains thereof, or encoding nucleic acids are also provided herein. Methods of expressing the encoded MTSP9 polypeptide and portions thereof using the cells are also provided, as are cells that express

-12-

MTSP9 on the cell surface. Such cells are used in methods of identifying candidate therapeutic compounds.

Additionally provided herein are antibodies that specifically bind to single and two chains forms of MTSP9, cells, combinations, kits and articles of manufacture that contain the antibodies. Antibodies that specifically bind to the MTSP9, particularly the single-chain protease domain, the two-chain form of the protease domain, the zymogen and activated form of MTSP9 and other fragments thereof. Neutralizing antibodies that inhibit a biological activity, particularly protease activity are also provided.

Further provided herein are prognostic, diagnostic, therapeutic screening methods using MTSP9 and the nucleic acids encoding MTSP9. In particular, the prognostic, diagnostic and therapeutic screening methods are used for preventing, treating, or for finding agents useful in preventing or treating, tumors or cancers such as lung carcinoma, colon adenocarcinoma and ovarian carcinoma.

Also provided herein are modulators of the activity of MTSP9, especially the modulators obtained according to the screening methods provide herein. Such modulators can have use in treating cancerous conditions and other neoplastic conditions.

Methods of diagnosing a disease or disorder characterized by detecting an aberrant level of an MTSP9 in a subject is provided. The method can be practiced by measuring the level of the DNA, RNA, protein or functional activity of the MTSP9. An increase or decrease in the level of the DNA, RNA, protein or functional activity of the MTSP, relative to the level of the DNA, RNA, protein or functional activity found in an analogous sample not having the disease or disorder (or other suitable control) is indicative of the presence of the disease or disorder in the subject or other relative any other suitable control.

Also provided are methods of identifying a compound that binds to the single-chain and/or two-chain form of MTSP9, by contacting a test compound with a both forms; determining to which form the compound binds; and if it binds to a form of MTSP9, further determining whether the compound has at least one of the following properties:

-13-

- (i) inhibits activation of the single-chain zymogen form of MTSP9;
- (ii) inhibits activity of the two-chain or single-chain form; and
- (iii) inhibits dimerization of the protein.

The forms can be full length or truncated forms, including but not limited to, the
5 protease domain resulting from cleavage at the activation cleavage site (between
amino acids R₁₈₅ and I₁₈₆); or from expression of the protease domain or
catalytically active portions thereof.

Pharmaceutical composition containing the protease domain and/or full-
length or other domain of an MTSP9 polypeptide are provided herein in a
10 pharmaceutically acceptable carrier or excipient are provided herein.

Also provided are articles of manufacture that contain MTSP9
polypeptide and protease domains of MTSP9 in single chain forms or activated
forms. The articles contain a) packaging material; b) the polypeptide (or
encoding nucleic acid), particularly the single chain protease domain thereof; and
15 c) a label indicating that the article is for using ins. assays for identifying
modulators of the activities of an MTSP9 polypeptide is provided herein.

Conjugates containing a) an MTSP9 polypeptide or protease domain in a
single or two chain form; and b) a targeting agent linked to the MTSP directly or
via a linker, wherein the agent facilitates: i) affinity isolation or purification of the
20 conjugate; ii) attachment of the conjugate to a surface; iii) detection of the
conjugate; or iv) targeted delivery to a selected tissue or cell, is provided herein.
The conjugate can contain a plurality of agents linked thereto. The conjugate
can be a chemical conjugate; and it can be a fusion protein. The targeting agent
can be a protein or peptide fragment. The protein or peptide fragment can
25 include a protein binding sequence, a nucleic acid binding sequence, a lipid
binding sequence, a polysaccharide binding sequence, or a metal binding
sequence.

Combinations are provided herein. The combination can include: a) an
inhibitor of the activity of an MTSP9; and b) an anti-cancer treatment or agent.
30 The MTSP inhibitor and the anti-cancer agent can be formulated in a single
pharmaceutical composition or each is formulated in a separate pharmaceutical
composition. The MTSP9 inhibitor can be an antibody or a fragment or binding

-14-

portion thereof made against the MTSP9, such as an antibody that specifically binds to the protease domain, an inhibitor of MTSP9 production, or an inhibitor of MTSP9 membrane-localization or an inhibitor of MTSP9 activation. Other MTSP9 inhibitors include, but are not limited to, an antisense nucleic acid or

5 double-stranded RNA (dsRNA), such as RNAi, encoding the MTSP9, particularly a portion of the protease domain; a nucleic acid encoding at least a portion of a gene encoding the MTSP9 with a heterologous nucleotide sequence inserted therein such that the heterologous sequence inactivates the biological activity encoded MTSP9 or the gene encoding it. For example, the portion of the gene

10 encoding the MTSP9 can flank the heterologous sequence to promote homologous recombination with a genomic gene encoding the MTSP9.

Also provided are methods for treating or preventing a tumor or cancer in a mammal by administering to a mammal an effective amount of an inhibitor of an MTSP9, whereby the tumor or cancer is treated or prevented. The MTSP9

15 inhibitor used in the treatment or for prophylaxis is administered with a pharmaceutically acceptable carrier or excipient. The mammal treated can be a human. The treatment or prevention method can additionally include administering an anti-cancer treatment or agent simultaneously with or subsequently or before administration of the MTSP9 inhibitor.

Also provided are transgenic non-human animals bearing inactivated genes encoding the MTSP and bearing the genes encoding the MTSP9 under non-native promotor control are provided. Such animals are useful in animal models of tumor initiation, growth and/or progression models. Transgenic non-human animals containing heterologous nucleic acid MTSP9 under native, non-

25 native promotor control or on an exogenous element, such as a plasmid or artificial chromosome, are additionally provided herein. In particular, recombinant non-human animals are provided herein, where the gene of an MTSP9 is under control of a promoter that is not the native promoter of the gene or that is not the native promoter of the gene in the non-human animal or where

30 the nucleic acid encoding the MTSP9 is heterologous to the non-human animal and the promoter is the native or a non-native promoter or the MTSP9 is on an extrachromosomal element, such as a plasmid or artificial chromosome.

-15-

Recombinant and transgenic animals can be produced by homologous recombination and non-homologous recombination methods.

Methods of gene therapy are provided. Such methods can be effected administering *in vivo* or *ex vivo* an inactivating form of the MTSP9 or by
5 administering an MTSP-encoding nucleic acid molecule are also provided.

Also provided are methods of treatments of tumors by administering a prodrug that is activated by MTSP9 that is expressed or active in tumor cells, particularly those in which its functional activity in tumor cells is greater than in non-tumor cells. The prodrug is administered and, upon administration, active
10 MTSP9 expressed on cells cleaves the prodrug and releases active drug in the vicinity of the tumor cells. The active anti-cancer drug accumulates in the vicinity of the tumor. This is particularly useful in instances in which MTSP9 is expressed or active in greater quantity, higher level or predominantly in tumor cells compared to other cells.

Also provided are methods of diagnosing the presence of a pre-malignant lesion, a malignancy, or other pathologic condition in a subject, by obtaining a biological sample from the subject; exposing it to a detectable agent that binds to a two-chain and/or single-chain form of MTSP9, where the pathological condition is characterized by the presence or absence of the two-chain and/or
15 20 single-chain form.

Methods of inhibiting tumor invasion or metastasis or treating a malignant or pre-malignant condition by administering an agent that inhibits activation of the zymogen form of MTSP9 or an activity of the activated form are provided. The conditions include, but are not limited to, a condition, such as a tumor, of
25 the breast, cervix, prostate, lung, ovary or colon.

DETAILED DESCRIPTION

A. DEFINITIONS

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as is commonly understood by one of skill in the art to
30 which the invention(s) belong. All patents, patent applications, published applications and publications, Genbank sequences, websites and other published materials referred to throughout the entire disclosure herein, unless noted

-16-

otherwise, are incorporated by reference in their entirety. In the event that there are a plurality of definitions for terms herein, those in this section prevail.

Where reference is made to a URL or other such identifier or address, it is understood that such identifiers can change and particular information on the internet can come and go, but equivalent information can be found by searching the internet. Reference thereto evidences the availability and public dissemination of such information.

As used herein, the abbreviations for any protective groups, amino acids and other compounds, are, unless indicated otherwise, in accord with their common usage, recognized abbreviations, or the IUPAC-IUB Commission on Biochemical Nomenclature (see, (1972) *Biochem.* 11:942-944).

As used herein, serine protease refers to a diverse family of proteases wherein a serine residue is involved in the hydrolysis of proteins or peptides. The serine residue can be part of the catalytic triad mechanism, which includes a serine, a histidine and an aspartic acid in the catalysis, or be part of the hydroxyl/ ϵ -amine or hydroxyl/ α -amine catalytic dyad mechanism, which involves a serine and a lysine in the catalysis. Of particular interest are SPs of mammalian, including human, origin. Those of skill in this art recognize that, in general, single amino acid substitutions in non-essential regions of a polypeptide do not substantially alter biological activity (see, *e.g.*, Watson *et al.* (1987) *Molecular Biology of the Gene*, 4th Edition, The Benjamin/Cummings Pub. co., p.224).

As used herein, "transmembrane serine protease (MTSP)" refers to a family of transmembrane serine proteases that share common structural features as described herein (see, also Hooper *et al.* (2001) *J. Biol. Chem.* 276:857-860). Thus, reference, for example, to "MTSP" encompasses all proteins encoded by the MTSP gene family, including but are not limited to: MTSP3, MTSP4, MTSP6, MTSP7 or an equivalent molecule obtained from any other source or that has been prepared synthetically or that exhibits the same activity. Other MTSPs include, but are not limited to, corin, enterpeptidase, human airway trypsin-like protease (HAT), MTSP1, TMPRSS2 and TMPRSS4. Sequences of encoding nucleic acid molecules and the encoded amino acid sequences of

-17-

exemplary MTSPs and/or domains thereof are set forth, for example in U.S. application Serial No. 09/776,191 (SEQ ID Nos. 1-12, 49, 50 and 61-72 therein, published as International PCT application No. WO 01/57194). The term also encompass MTSPs with amino acid substitutions that do not substantially alter activity of each member and also encompasses splice variants thereof. Suitable substitutions, including, although not necessarily, conservative substitutions of amino acids, are known to those of skill in this art and can be made without eliminating the biological activity, such as the catalytic activity, of the resulting molecule.

10 As used herein an MTSP9, whenever referenced herein, includes at least one or all of or any combination of:

a polypeptide encoded by the sequence of nucleotides set forth in SEQ ID No. 17 or by a sequence of nucleotides that includes nucleotides that encode amino acids 11-232 of SEQ ID No. 6;

15 a polypeptide encoded by a sequence of nucleotides that hybridizes under conditions of low, moderate or high stringency to the sequence of nucleotides set forth in is set forth as nucleotides 31-729 SEQ ID No. 5 or as SEQ ID No. 17;

a polypeptide that includes the sequence of amino acids set forth as amino acids 11-232 of SEQ ID No. 6;

20 a polypeptide that includes a sequence of amino acids having at least about 60%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity with the sequence of amino acids set forth in SEQ ID No. 17 or 18 or as amino acids 11-232 of SEQ ID No. 6; and/or

25 a polypeptide encoded by a splice variant of the MTSP9 set forth in SEQ ID No. 17.

In particular, the MTSP9 polypeptide, with the protease domains as indicated in SEQ ID No. 5, 6, 16, 17 and 18, is provided. The polypeptide is a single or two chain polypeptide. Smaller portions thereof that retain protease activity are also provided. The protease domains from MTSPs vary in size and constitution, including insertions and deletions in surface loops. They retain

30

-18-

conserved structure, including at least one of the active site triad (see, *e.g.*, the catalytic triad of the MTSP9 in SEQ ID No. 18 is H₂₂₇, D₂₇₂ and S₃₆₈), primary specificity pocket, oxyanion hole and/or other features of serine protease domains of proteases. Thus, for purposes herein, the protease domain is a
5 portion of an MTSP, as defined herein, and is homologous to a domain of other MTSPs, such as corin, enterpeptidase, human airway trypsin-like protease (HAT), MTSP1, TMPRSS2, and TMPRSS4, which have been previously identified; it was not recognized, however, that an isolated single chain form of the protease domain could function proteolytically in *in vitro* assays. As with
10 the larger class of enzymes of the chymotrypsin (S1) fold (see, *e.g.*, Internet accessible MEROPS data base), the MTSPs protease domains share a high degree of amino acid sequence identity. The His, Asp and Ser residues necessary for activity are present in conserved motifs. The activation site, which results in the N-terminus of the second chain in the two chain form is
15 located in a conserved motif and readily can be identified.

The MTSP9 can be from any animal, particularly a mammal, and includes but are not limited to, humans, rodents, fowl, ruminants and other animals. The full-length zymogen or two-chain activated form is contemplated or any domain thereof, including the protease domain, which can be a two-chain activated
20 form, or a single chain form.

As used herein, a "protease domain of an MTSP" refers to an extracellular protease domain of an MTSP that exhibits proteolytic activity and shares homology and structural features with the chymotrypsin/trypsin family protease domains. Hence it is at least the minimal portion of the domain that exhibits
25 proteolytic activity as assessed by standard *in vitro* assays. Contemplated herein are such protease domains and catalytically active portions thereof. Also provided are truncated forms of the protease domain that include the smallest fragment thereof that acts catalytically as a single chain form.

A protease domain of an MTSP9, whenever referenced herein, includes at
30 least one or all of or any combination of or a catalytically active portion of:
a polypeptide that includes the sequence of amino acids set forth in SEQ ID No. 16 ;

-19-

a polypeptide encoded by a sequence of nucleotides that hybridizes under conditions of low, moderate or high stringency to the sequence of nucleotides set forth in SEQ ID No. 15 or 17;

a polypeptide that includes the sequence of amino acids set forth
5 in SEQ ID No. 6, 16 or 18;

a polypeptide that includes a sequence of amino acids having at least about 60%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity with the sequence of amino acids set forth in SEQ ID No. 6,
10 16; or 18; and/or

a protease domain of a polypeptide encoded by a splice variant of the MTSP9.

The protease domains of MTSPs vary in size and constitution, including insertions and deletions in surface loops. They retain conserved structure,
15 including at least one of the active site triad, primary specificity pocket, oxyanion hole and/or other features of serine protease domains of proteases. Thus, for purposes herein, the protease domain is a portion of an MTSP, as defined herein, and is homologous to a domain of other MTSP. As with the larger class of enzymes of the chymotrypsin (S1) fold (see, *e.g.*, Internet
20 accessible MEROPS data base), the MTSP protease domains share a high degree of amino acid sequence identity. The His, Asp and Ser residues necessary for activity are present in conserved motifs. The activation site, whose cleavage creates the N-terminus of the protease domain in the two-chain forms is located in a conserved motif and readily can be identified.

25 By active form is meant a form active *in vivo* and/or *in vitro*. As described herein, the protease domain also can exist as a two-chain form. It is shown herein that, at least *in vitro*, the single chain forms of the SPs and the catalytic domains or proteolytically active portions thereof (typically C-terminal truncations) exhibit protease activity. Hence provided herein are isolated single
30 chain forms of the protease domains of SPs and their use in *in vitro* drug screening assays for identification of agents that modulate the activity thereof.

-20-

As used herein, the catalytically active domain of an MTSP refers to the protease domain. Reference to the protease domain of an MTSP generally refers to the single chain form of the protein. If the two-chain form or both forms is intended, it is so-specified. The zymogen form of each protein is a single chain,
5 which is converted to the active two chain form by activation cleavage.

As used herein, activation cleavage refers to the cleavage of the protease at the N-terminus of the protease domain (in this instance between R₁₈₅ and I₁₈₆; with reference to SEQ ID Nos. 12 and 13). By virtue of the Cys-Cys pairing between a Cys outside the protease domain (in this instance, for example C₁₇₅)
10 and a Cys in the protease domain (in this instance Cys₂₉₂) upon cleavage the resulting polypeptide has two chains ("A" chain and the "B" chain, which is the protease domain). Cleavage can be effected by another protease or autocatalytically.

As used herein, a two-chain form of the protease domain refers to a two-
15 chain form that is formed from the two-chain form of the protease in which the Cys pairing between, in this instance, Cys₁₇₅ and Cys₂₉₂, which links the protease domain to the remainder of the polypeptide, the "A" chain. A two chain protease domain form refers to any form in which the "remainder of the polypeptide", *i.e.*, "A" chain, is shortened and includes from at Cys₁₇₅.

20 MTSPs of interest include those that are activated and/or expressed in tumor cells different, typically higher, from those in non-tumor cells; and those from cells in which substrates therefor differ in tumor cells from non-tumor cells or differ with respect to substrates, co-factors or receptors, or otherwise alter the activity or specificity of the MTSP.

25 ^F^F As used herein, a human protein is one encoded by nucleic acid, such as DNA, present in the genome of a human, including all allelic variants and conservative variations as long as they are not variants found in other mammals.

As used herein, a "nucleic acid encoding a protease domain or catalytically active portion of a SP" shall be construed as referring to a nucleic
30 acid encoding only the recited single chain protease domain or active portion thereof, and not the other contiguous portions of the SP as a continuous sequence.

-21-

As used herein, catalytic activity refers to the activity of the SP as a serine protease. Function of the SP refers to its function in tumor biology, including promotion of or involvement in initiation, growth or progression of tumors, and also roles in signal transduction. Catalytic activity refers to the activity of the SP as a protease as assessed in *in vitro* proteolytic assays that
5 detect proteolysis of a selected substrate.

As used herein, a zymogen is an inactive precursor of a proteolytic enzyme. Such precursors are generally larger, although not necessarily larger than the active form. With reference to serine proteases, zymogens are
10 converted to active enzymes by specific cleavage, including catalytic and autocatalytic cleavage, or by binding of an activating co-factor, which generates the an active enzyme. A zymogen, thus, is an enzymatically inactive protein that is converted to a proteolytic enzyme by the action of an activator.

As used herein, "disease or disorder" refers to a pathological condition in
15 an organism resulting from, *e.g.*, infection or genetic defect, and characterized by identifiable symptoms.

As used herein, neoplasm (neoplasia) refers to abnormal new growth, and thus means the same as *tumor*, which can be benign or malignant. Unlike *hyperplasia*, neoplastic proliferation persists even in the absence of the original
20 stimulus.

As used herein, neoplastic disease refers to any disorder involving cancer, including tumor development, growth, metastasis and progression.

As used herein, cancer refers to a general term for diseases caused by any type of malignant tumor.

25 As used herein, malignant, as applies to tumors, refers to primary tumors that have the capacity of *metastasis* with loss of *growth control* and *positional control*.

As used herein, an anti-cancer agent (used interchangeable with "anti-tumor or anti-neoplastic agent") refers to any agents used in the anti-cancer
30 treatment. These include any agents, when used alone or in combination with other compounds, that can alleviate, reduce, ameliorate, prevent, or place or maintain in a state of remission of clinical symptoms or diagnostic markers

-22-

associated with neoplastic disease, tumor and cancer, and can be used in methods, combinations and compositions provided herein. Non-limiting examples of anti-neoplastic agents include anti-angiogenic agents, alkylating agents, antimetabolites, certain natural products, platinum coordination
5 complexes, anthracenediones, substituted ureas, methylhydrazine derivatives, adrenocortical suppressants, certain hormones, antagonists and anti-cancer polysaccharides.

As used herein, a splice variant refers to a variant produced by differential processing of a primary transcript of genomic nucleic acid, such as DNA, that
10 results in more than one type of mRNA. Splice variants of SPs are provided herein.

As used herein, angiogenesis is intended to broadly encompass the totality of processes directly or indirectly involved in the establishment and maintenance of new vasculature (neovascularization), including, but not limited
15 to, neovascularization associated with tumors.

As used herein, anti-angiogenic treatment or agent refers to any therapeutic regimen and compound, when used alone or in combination with other treatment or compounds, that can alleviate, reduce, ameliorate, prevent, or place or maintain in a state of remission of clinical symptoms or diagnostic
20 markers associated with undesired and/or uncontrolled angiogenesis. Thus, for purposes herein an anti-angiogenic agent refers to an agent that inhibits the establishment or maintenance of vasculature. Such agents include, but are not limited to, anti-tumor agents, and agents for treatments of other disorders associated with undesirable angiogenesis, such as diabetic retinopathies,
25 restenosis, hyperproliferative disorders and others.

As used herein, non-anti-angiogenic anti-tumor agents refer to anti-tumor agents that do not act primarily by inhibiting angiogenesis.

As used herein, pro-angiogenic agents are agents that promote the establishment or maintenance of the vasculature. Such agents include agents
30 for treating cardiovascular disorders, including heart attacks and strokes.

As used herein, undesired and/or uncontrolled angiogenesis refers to pathological angiogenesis wherein the influence of angiogenesis stimulators

-23-

outweighs the influence of angiogenesis inhibitors. As used herein, deficient angiogenesis refers to pathological angiogenesis associated with disorders where there is a defect in normal angiogenesis resulting in aberrant angiogenesis or an absence or substantial reduction in angiogenesis.

5 As used herein, the protease domain of an SP protein refers to the protease domain of an SP that exhibits proteolytic activity. Hence it is at least the minimal portion of the protein that exhibits proteolytic activity as assessed by standard assays *in vitro*. It refers, herein, to a single chain form and also the two chain activated form (where the two chain form is intended it will be so-
10 noted). Exemplary protease domains include at least a sufficient portion of sequences of amino acids set forth in SEQ ID No. 6 (encoded by nucleotides in SEQ ID No. 5) to exhibit protease activity.

 Also contemplated are nucleic acid molecules that encode a polypeptide that has proteolytic activity in an *in vitro* proteolysis assay and that have at least
15 60%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity with the full-length of a protease domain of an MTSP9 polypeptide, or that hybridize along their full-length or along at least about 70%, 80% or 90% of the full-length to a nucleic acids that encode a protease domain, particularly
20 under conditions of moderate, generally high, stringency.

 For the protease domains, residues at the N-terminus can be critical for activity. It is shown herein that the protease domain of the single chain form of the MTSP9 protease is catalytically active. Hence the protease domain generally requires the N-terminal amino acids thereof for activity; the C-terminus portion
25 can be truncated. The amount that can be removed can be determined empirically by testing the polypeptide for protease activity in an *in vitro* assay that assesses catalytic cleavage.

 Hence smaller portions of the protease domains, particularly the single chain domains, thereof that retain protease activity are contemplated. Such
30 smaller versions generally are C-terminal truncated versions of the protease domains. The protease domains vary in size and constitution, including insertions and deletions in surface loops. Such domains exhibit conserved

-24-

structure, including at least one structural feature, such as the active site triad, primary specificity pocket, oxyanion hole and/or other features of serine protease domains of proteases. Thus, for purposes herein, the protease domain is a single chain portion of an MTSP9, as defined herein, but is homologous in its structural features and retention of sequence of similarity or homology the protease domain of chymotrypsin or trypsin. The polypeptide exhibits proteolytic activity as a single chain.

As used herein, by homologous means about greater than 25% nucleic acid sequence identity, such as 25%, 40%, 60%, 70%, 80%, 90% or 95%. If necessary the percentage homology will be specified. The terms "homology" and "identity" are often used interchangeably. In general, sequences are aligned so that the highest order match is obtained (see, e.g.: *Computational Molecular Biology*, Lesk, A.M., ed., Oxford University Press, New York, 1988;

Biocomputing: Informatics and Genome Projects, Smith, D.W., ed., Academic Press, New York, 1993; *Computer Analysis of Sequence Data, Part I*, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994; *Sequence Analysis in Molecular Biology*, von Heinje, G., Academic Press, 1987; and *Sequence Analysis Primer*, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991; Carillo *et al.* (1988) *SIAM J Applied Math* 48:1073).

By sequence identity, the number of conserved amino acids are determined by standard alignment algorithms programs, and are used with default gap penalties established by each supplier. Substantially homologous nucleic acid molecules would hybridize typically at moderate stringency or at high stringency all along the length of the nucleic acid or or along at least about 70%, 80% or 90% of the full-length nucleic acid molecule of interest. Also contemplated are nucleic acid molecules that contain degenerate codons in place of codons in the hybridizing nucleic acid molecule.

Whether any two nucleic acid molecules have nucleotide sequences that are at least, for example, 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99%

"identical" can be determined using known computer algorithms such as the "FAST A" program, using for example, the default parameters as in Pearson *et al.* (1988) *Proc. Natl. Acad. Sci. USA* 85:2444 (other programs include the GCG

-25-

program package (Devereux, J., *et al.*, *Nucleic Acids Research* 12(1):387 (1984)), BLASTP, BLASTN, FASTA (Atschul, S.F., *et al.*, *J Molec Biol* 215:403 (1990); Guide to Huge Computers, Martin J. Bishop, ed., Academic Press, San Diego, 1994, and Carillo *et al.* (1988) *SIAM J Applied Math* 48:1073). For
5 example, the BLAST function of the National Center for Biotechnology Information database can be used to determine identity. Other commercially or publicly available programs include, DNASTar "MegAlign" program (Madison, WI) and the University of Wisconsin Genetics Computer Group (UWG) "Gap" program (Madison WI). Percent homology or identity of proteins and/or nucleic
10 acid molecules can be determined, for example, by comparing sequence information using a GAP computer program (*e.g.*, Needleman *et al.* (1970) *J. Mol. Biol.* 48:443, as revised by Smith and Waterman ((1981) *Adv. Appl. Math.* 2:482). Briefly, the GAP program defines similarity as the number of aligned symbols (*i.e.*, nucleotides or amino acids) which are similar, divided by the total
15 number of symbols in the shorter of the two sequences. Default parameters for the GAP program can include: (1) a unary comparison matrix (containing a value of 1 for identities and 0 for non-identities) and the weighted comparison matrix of Gribskov *et al.* (1986) *Nucl. Acids Res.* 14:6745, as described by Schwartz and Dayhoff, eds., *ATLAS OF PROTEIN SEQUENCE AND STRUCTURE*, National
20 Biomedical Research Foundation, pp. 353-358 (1979); (2) a penalty of 3.0 for each gap and an additional 0.10 penalty for each symbol in each gap; and (3) no penalty for end gaps. Therefore, as used herein, the term "identity" represents a comparison between a test and a reference polypeptide or polynucleotide.

As used herein, the term at least "90% identical to" refers to percent
25 identities from 90 to 99.99 relative to the reference polypeptides. Identity at a level of 90% or more is indicative of the fact that, assuming for exemplification purposes a test and reference polynucleotide length of 100 amino acids are compared. No more than 10% (*i.e.*, 10 out of 100) amino acids in the test polypeptide differs from that of the reference polypeptides. Similar comparisons
30 can be made between a test and reference polynucleotides. Such differences can be represented as point mutations randomly distributed over the entire length of an amino acid sequence or they can be clustered in one or more

-26-

locations of varying length up to the maximum allowable, e.g. 10/100 amino acid difference (approximately 90% identity). Differences are defined as nucleic acid or amino acid substitutions, or deletions. At the level of homologies or identities above about 85-90%, the result should be independent of the program and gap parameters set; such high levels of identity can be assessed readily, often without relying on software.

As used herein, primer refers to an oligonucleotide containing two or more deoxyribonucleotides or ribonucleotides, typically more than three, from which synthesis of a primer extension product can be initiated. Experimental conditions conducive to synthesis include the presence of nucleoside triphosphates and an agent for polymerization and extension, such as DNA polymerase, and a suitable buffer, temperature and pH.

As used herein, animals include any animal, such as, but are not limited to, goats, cows, deer, sheep, rodents, pigs and humans. Non-human animals, exclude humans as the contemplated animal. The SPs provided herein are from any source, animal, plant, prokaryotic and fungal. Most MTSP9s are of animal origin, including mammalian origin.

As used herein, genetic therapy involves the transfer of heterologous nucleic acid, such as DNA, into certain cells, target cells, of a mammal, particularly a human, with a disorder or conditions for which such therapy is sought. The nucleic acid, such as DNA, is introduced into the selected target cells in a manner such that the heterologous nucleic acid, such as DNA, is expressed and a therapeutic product encoded thereby is produced. Alternatively, the heterologous nucleic acid, such as DNA, can in some manner mediate expression of DNA that encodes the therapeutic product, or it can encode a product, such as a peptide or RNA that in some manner mediates, directly or indirectly, expression of a therapeutic product. Genetic therapy can also be used to deliver nucleic acid encoding a gene product that replaces a defective gene or supplements a gene product produced by the mammal or the cell in which it is introduced. The introduced nucleic acid can encode a therapeutic compound, such as a growth factor inhibitor thereof, or a tumor necrosis factor or inhibitor thereof, such as a receptor therefor, that is not

-27-

normally produced in the mammalian host or that is not produced in therapeutically effective amounts or at a therapeutically useful time. The heterologous nucleic acid, such as DNA, encoding the therapeutic product can be modified prior to introduction into the cells of the afflicted host in order to
5 enhance or otherwise alter the product or expression thereof. Genetic therapy can also involve delivery of an inhibitor or repressor or other modulator of gene expression.

As used herein, heterologous nucleic acid is nucleic acid that (if DNA encodes RNA) and proteins that are not normally produced *in vivo* by the cell in
10 which it is expressed or that mediates or encodes mediators that alter expression of endogenous nucleic acid, such as DNA, by affecting transcription, translation, or other regulatable biochemical processes. Heterologous nucleic acid, such as DNA, can also be referred to as foreign nucleic acid, such as DNA. Any nucleic acid, such as DNA, that one of skill in the art would recognize or consider as
15 heterologous or foreign to the cell in which is expressed is herein encompassed by heterologous nucleic acid; heterologous nucleic acid includes exogenously added nucleic acid that is also expressed endogenously. Examples of heterologous nucleic acid include, but are not limited to, nucleic acid that encodes traceable marker proteins, such as a protein that confers drug
20 resistance, nucleic acid that encodes therapeutically effective substances, such as anti-cancer agents, enzymes and hormones, and nucleic acid, such as DNA, that encodes other types of proteins, such as antibodies. Antibodies that are encoded by heterologous nucleic acid can be secreted or expressed on the surface of the cell in which the heterologous nucleic acid has been introduced.
25 Heterologous nucleic acid is generally not endogenous to the cell into which it is introduced, but has been obtained from another cell or prepared synthetically. Generally, although not necessarily, such nucleic acid encodes RNA and proteins that are not normally produced by the cell in which it is expressed.

As used herein, a therapeutically effective product is a product that is
30 encoded by heterologous nucleic acid, typically DNA, that, upon introduction of the nucleic acid into a host, a product is expressed that ameliorates or eliminates

-28-

the symptoms, manifestations of an inherited or acquired disease or that cures the disease.

As used herein, recitation that a polypeptide consists essentially of the protease domain means that the only SP portion of the polypeptide is a protease domain or a catalytically active portion thereof. The polypeptide can optionally,
5 and generally will, include additional non-SP-derived sequences of amino acids.

As used herein, cancer or tumor treatment or agent refers to any therapeutic regimen and/or compound that, when used alone or in combination with other treatments or compounds, can alleviate, reduce, ameliorate, prevent,
10 or place or maintain in a state of remission of clinical symptoms or diagnostic markers associated with deficient angiogenesis.

As used herein, domain refers to a portion of a molecule, *e.g.*, proteins or the encoding nucleic acids, that is structurally and/or functionally distinct from other portions of the molecule.

As used herein, protease refers to an enzyme catalyzing hydrolysis of proteins or peptides. It includes the zymogen form and activated forms thereof. For clarity reference to protease refers to all forms, and particular forms will be specifically designated. For purposes herein, the protease domain includes single and two chain forms of the protease domain of an SP protein. For MTSP9 the
15 protease domain also includes single and two chain forms of the protease domain.

As used herein, nucleic acids include DNA, RNA and analogs thereof, including protein nucleic acids (PNA) and mixture thereof. Nucleic acids can be single or double-stranded. When referring to probes or primers, optionally
25 labeled, with a detectable label, such as a fluorescent or radiolabel, single-stranded molecules are contemplated. Such molecules are typically of a length such that their target is statistically unique or of low copy number (typically less than 5, generally less than 3) for probing or priming a library. Generally a probe or primer contains at least 14, 16 or 30 contiguous of sequence complementary
30 to or identical a gene of interest. Probes and primers can be 10, 20, 30, 50, 100 or more nucleic acids long.

As used herein, nucleic acid encoding a fragment or portion of an SP refers to a nucleic acid encoding only the recited fragment or portion of SP, and not the other contiguous portions of the SP.

As used herein, operative linkage of heterologous nucleic to regulatory
5 and effector sequences of nucleotides, such as promoters, enhancers,
transcriptional and translational stop sites, and other signal sequences refers to
the relationship between such nucleic acid, such as DNA, and such sequences of
nucleotides. For example, operative linkage of heterologous DNA to a promoter
refers to the physical relationship between the DNA and the promoter such that
10 the transcription of such DNA is initiated from the promoter by an RNA
polymerase that specifically recognizes, binds to and transcribes the DNA in
reading frame. Thus, operatively linked or operationally associated refers to the
functional relationship of nucleic acid, such as DNA, with regulatory and effector
sequences of nucleotides, such as promoters, enhancers, transcriptional and
15 translational stop sites, and other signal sequences. For example, operative
linkage of DNA to a promoter refers to the physical and functional relationship
between the DNA and the promoter such that the transcription of such DNA is
initiated from the promoter by an RNA polymerase that specifically recognizes,
binds to and transcribes the DNA. In order to optimize expression and/or *in vitro*
20 transcription, it can be necessary to remove, add or alter 5' untranslated portions
of the clones to eliminate extra, potential inappropriate alternative translation
initiation (*i.e.*, start) codons or other sequences that can interfere with or reduce
expression, either at the level of transcription or translation. Alternatively,
consensus ribosome binding sites (see, *e.g.*, Kozak *J. Biol. Chem.* 266:19867-
25 19870 (1991)) can be inserted immediately 5' of the start codon and can
enhance expression. The desirability of (or need for) such modification can be
empirically determined.

As used herein, a sequence complementary to at least a portion of an
RNA, with reference to antisense oligonucleotides, means a sequence having
30 sufficient complementarity to be able to hybridize with the RNA, generally under
moderate or high stringency conditions, forming a stable duplex; in the case of
double-stranded SP antisense nucleic acids, a single strand of the duplex DNA

-30-

(or dsRNA) can thus be tested, or triplex formation can be assayed. The ability to hybridize depends on the degree of complementarity and the length of the antisense nucleic acid. Generally, the longer the hybridizing nucleic acid, the more base mismatches with a SP encoding RNA it can contain and still form a stable duplex (or triplex, as the case can be). One skilled in the art can ascertain a tolerable degree of mismatch by use of standard procedures to determine the melting point of the hybridized complex.

For purposes herein, amino acid substitutions can be made in any of SPs and protease domains thereof provided that the resulting protein exhibits protease activity. Amino acid substitutions contemplated include conservative substitutions, such as those set forth in Table 1, which do not eliminate proteolytic activity. As described herein, substitutions that alter properties of the proteins, such as removal of cleavage sites and other such sites are also contemplated; such substitutions are generally non-conservative, but can be readily effected by those of skill in the art.

Suitable conservative substitutions of amino acids are known to those of skill in this art and can be made generally without altering the biological activity, for example enzymatic activity, of the resulting molecule. Those of skill in this art recognize that, in general, single amino acid substitutions in non-essential regions of a polypeptide do not substantially alter biological activity (see, *e.g.*, Watson *et al. Molecular Biology of the Gene*, 4th Edition, 1987, The Bejacmin/Cummings Pub. co., p.224).• Also included within the definition, is the catalytically active fragment of an SP, particularly a single chain protease portion. Conservative amino acid substitutions are made, for example, in accordance with those set forth in TABLE 1 as follows:

-31-

TABLE 1

	Original residue	Conservative substitution
	Ala (A)	Gly; Ser, Abu
	Arg (R)	Lys, orn
5	Asn (N)	Gln; His
	Cys (C)	Ser
	Gln (Q)	Asn
	Glu (E)	Asp
	Gly (G)	Ala; Pro
10	His (H)	Asn; Gln
	Ile (I)	Leu; Val; Met; Nle; Nva
	Leu (L)	Ile; Val; Met; Nle; Nv
	Lys (K)	Arg; Gln; Glu
	Met (M)	Leu; Tyr; Ile; NLe Val
15	Ornithine	Lys; Arg
	Phe (F)	Met; Leu; Tyr
	Ser (S)	Thr
	Thr (T)	Ser
	Trp (W)	Tyr
20	Tyr (Y)	Trp; Phe
	Val (V)	Ile; Leu; Met; Nle; Nv

Other substitutions are also permissible and can be determined empirically or in accord with known conservative substitutions.

As used herein, Abu is 2-aminobutyric acid; Orn is ornithine.

25 As used herein, the amino acids, which occur in the various amino acid sequences appearing herein, are identified according to their well-known, three-letter or one-letter abbreviations. The nucleotides, which occur in the various DNA fragments, are designated with the standard single-letter designations used routinely in the art.

30 As used herein, a probe or primer based on a nucleotide sequence disclosed herein, includes at least 10, 14, typically at least 16 contiguous sequence of nucleotides of SEQ ID No. 5, and probes of at least 30, 50 or 100 contiguous sequence of nucleotides of SEQ ID No. 5. The length of the probe or primer for unique hybridization is a function of the complexity of the genome of
35 interest.

As used herein, amelioration of the symptoms of a particular disorder by administration of a particular pharmaceutical composition refers to any lessening, whether permanent or temporary, lasting or transient that can be attributed to or associated with administration of the composition.

-32-

As used herein, antisense polynucleotides refer to synthetic sequences of nucleotide bases complementary to mRNA or the sense strand of double-stranded DNA. Admixture of sense and antisense polynucleotides under appropriate conditions leads to the binding of the two molecules, or

5 hybridization. When these polynucleotides bind to (hybridize with) mRNA, inhibition of protein synthesis (translation) occurs. When these polynucleotides bind to double-stranded DNA, inhibition of RNA synthesis (transcription) occurs. The resulting inhibition of translation and/or transcription leads to an inhibition of the synthesis of the protein encoded by the sense strand. Antisense nucleic

10 acid molecule typically contain a sufficient number of nucleotides to specifically bind to a target nucleic acid, generally at least 5 contiguous nucleotides, often at least 14 or 16 or 30 contiguous nucleotides or modified nucleotides complementary to the coding portion of a nucleic acid molecule that encodes a gene of interest, for example, nucleic acid encoding a single chain protease

15 domain of an SP.

As used herein, an array refers to a collection of elements, such as antibodies, containing three or more members. An addressable array is one in which the members of the array are identifiable, typically by position on a solid phase support. Hence, in general the members of the array are immobilized on

20 discrete identifiable loci on the surface of a solid phase.

As used herein, antibody refers to an immunoglobulin, whether natural or partially or wholly synthetically produced, including any derivative thereof that retains the specific binding ability the antibody. Hence antibody includes any protein having a binding domain that is homologous or substantially homologous

25 to an immunoglobulin binding domain. Antibodies include members of any immunoglobulin claims, including IgG, IgM, IgA, IgD and IgE.

As used herein, antibody fragment refers to any derivative of an antibody that is less than full-length, retaining at least a portion of the full-length antibody's specific binding ability. Examples of antibody fragments include, but

30 are not limited to, Fab, Fab', F(ab)₂, single-chain Fvs (scFV), FV, dsFV diabody and Fd fragments. The fragment can include multiple chains linked together,

-33-

such as by disulfide bridges. An antibody fragment generally contains at least about 50 amino acids and typically at least 200 amino acids.

As used herein, an Fv antibody fragment is composed of one variable heavy domain (V_H) and one variable light domain linked by noncovalent
5 interactions.

As used herein, a dsFV refers to an Fv with an engineered intermolecular disulfide bond, which stabilizes the V_H - V_L pair.

As used herein, an $F(ab)_2$ fragment is an antibody fragment that results from digestion of an immunoglobulin with pepsin at pH 4.0-4.5; it can be
10 recombinantly expressed to produce the equivalent fragment.

As used herein, Fab fragments are antibody fragments that result from digestion of an immunoglobulin with papain; they can be recombinantly expressed to produce the equivalent fragment.

As used herein, scFVs refer to antibody fragments that contain a variable
15 light chain (V_L) and variable heavy chain (V_H) covalently connected by a polypeptide linker in any order. The linker is of a length such that the two variable domains are bridged without substantial interference. Included linkers are $(Gly-Ser)_n$ residues with some Glu or Lys residues dispersed throughout to increase solubility.

As used herein, humanized antibodies refer to antibodies that are
20 modified to include human sequences of amino acids so that administration to a human does not provoke an immune response. Methods for preparation of such antibodies are known. For example, to produce such antibodies, the hybridoma or other prokaryotic or eukaryotic cell, such as an *E. coli* or a CHO cell, that
25 expresses the monoclonal antibody are altered by recombinant DNA techniques to express an antibody in which the amino acid composition of the non-variable region is based on human antibodies. Computer programs have been designed to identify such regions.

As used herein, diabodies are dimeric scFV; diabodies typically have
30 shorter peptide linkers than scFVs, and they generally dimerize.

-34-

As used herein, production by recombinant means by using recombinant DNA methods means the use of the well known methods of molecular biology for expressing proteins encoded by cloned DNA.

As used herein the term assessing is intended to include

5 quantitative and qualitative determination in the sense of obtaining an absolute value for the activity of an SP, or a domain thereof, present in the sample, and also of obtaining an index, ratio, percentage, visual or other value indicative of the level of the activity. Assessment can be direct or indirect and the chemical species actually detected need not of course be the proteolysis

10 product itself but can for example be a derivative thereof or some further substance.

As used herein, biological activity refers to the *in vivo* activities of a compound or physiological responses that result upon *in vivo* administration of a compound, composition or other mixture. Biological activity, thus, encompasses

15 therapeutic effects and pharmaceutical activity of such compounds, compositions and mixtures. Biological activities can be observed in *in vitro* systems designed to test or use such activities. Thus, for purposes herein the biological activity of a luciferase is its oxygenase activity whereby, upon oxidation of a substrate, light is produced.

As used herein, functional activity refers to a polypeptide or portion thereof that displays one or more activities associated with a full-length protein. Functional activities include, but are not limited to, biological activity, catalytic or enzymatic activity, antigenicity (ability to bind to or compete with a polypeptide for binding to an anti-polypeptide antibody), immunogenicity, ability to form

20 multimers, the ability to specifically bind to a receptor or ligand for the polypeptide.

25

As used herein, a conjugate refers to the compounds provided herein that include one or more SPs, including an MTSP9, particularly single chain protease domains thereof, and one or more targeting agents. These conjugates include

30 those produced by recombinant means as fusion proteins, those produced by chemical means, such as by chemical coupling, through, for example, coupling to sulfhydryl groups, and those produced by any other method whereby at least

-35-

one SP, or a domain thereof, is linked, directly or indirectly via linker(s) to a targeting agent.

As used herein, a targeting agent is any moiety, such as a protein or effective portion thereof, that provides specific binding of the conjugate to a cell surface receptor, which, can internalize the conjugate or SP portion thereof. A
5 targeting agent can also be one that promotes or facilitates, for example, affinity isolation or purification of the conjugate; attachment of the conjugate to a surface; or detection of the conjugate or complexes containing the conjugate.

As used herein, an antibody conjugate refers to a conjugate in which the
10 targeting agent is an antibody.

As used herein, derivative or analog of a molecule refers to a portion derived from or a modified version of the molecule.

As used herein, an effective amount of a compound for treating a particular disease is an amount that is sufficient to ameliorate, or in some
15 manner reduce the symptoms associated with the disease. Such amount can be administered as a single dosage or can be administered according to a regimen, whereby it is effective. The amount can cure the disease but, typically, is administered in order to ameliorate the symptoms of the disease. Repeated administration can be required to achieve the desired amelioration of symptoms.

As used herein equivalent, when referring to two sequences of nucleic
20 acids means that the two sequences in question encode the same sequence of amino acids or equivalent proteins. When equivalent is used in referring to two proteins or peptides, it means that the two proteins or peptides have substantially the same amino acid sequence with only amino acid substitutions
25 (such, as but not limited to, conservative changes such as those set forth in Table 1, above) that do not substantially alter the activity or function of the protein or peptide. When equivalent refers to a property, the property does not need to be present to the same extent (*e.g.*, two peptides can exhibit different rates of the same type of enzymatic activity), but the activities are usually
30 substantially the same. Complementary, when referring to two nucleotide sequences, means that the two sequences of nucleotides are capable of hybridizing, typically with less than 25%, 15%, 5% or 0% mismatches between

-36-

opposed nucleotides. If necessary the percentage of complementarity will be specified. Typically the two molecules are selected such that they will hybridize under conditions of high stringency.

As used herein, an agent that modulates the activity of a protein or
5 expression of a gene or nucleic acid either decreases or increases or otherwise alters the activity of the protein or, in some manner up- or down-regulates or otherwise alters expression of the nucleic acid in a cell.

As used herein, inhibitor of the activity of an SP encompasses any substance that prohibits or decrease production, post-translational
10 modification(s), maturation, or membrane localization of the SP or any substance that interferes with or decreases the proteolytic efficacy of thereof, particularly of a single chain form in an *in vitro* screening assay.

As used herein, a method for treating or preventing neoplastic disease means that any of the symptoms, such as the tumor, metastasis thereof, the
15 vascularization of the tumors or other parameters by which the disease is characterized are reduced, ameliorated, prevented, placed in a state of remission, or maintained in a state of remission. It also means that the hallmarks of neoplastic disease and metastasis can be eliminated, reduced or prevented by the treatment. Non-limiting examples of the hallmarks include uncontrolled
20 degradation of the basement membrane and proximal extracellular matrix, migration, division, and organization of the endothelial cells into new functioning capillaries, and the persistence of such functioning capillaries.

As used herein, pharmaceutically acceptable salts, esters or other derivatives of the conjugates include any salts, esters or derivatives that can be
25 readily prepared by those of skill in this art using known methods for such derivatization and that produce compounds that can be administered to animals or humans without substantial toxic effects and that either are pharmaceutically active or are prodrugs.

As used herein, a prodrug is a compound that, upon *in vivo*
30 administration, is metabolized or otherwise converted to the biologically, pharmaceutically or therapeutically active form of the compound. To produce a prodrug, the pharmaceutically active compound is modified such that the active

-37-

compound is regenerated by metabolic processes. The prodrug can be designed to alter the metabolic stability or the transport characteristics of a drug, to mask side effects or toxicity, to improve the flavor of a drug or to alter other characteristics or properties of a drug. By virtue of knowledge of

5 pharmacodynamic processes and drug metabolism *in vivo*, those of skill in this art, once a pharmaceutically active compound is known, can design prodrugs of the compound (see, *e.g.*, Nogrady (1985) *Medicinal Chemistry A Biochemical Approach*, Oxford University Press, New York, pages 388-392).

As used herein, a drug identified by the screening methods provided
10 herein refers to any compound that is a candidate for use as a therapeutic or as a lead compound for the design of a therapeutic. Such compounds can be small molecules, including small organic molecules, peptides, peptide mimetics, antisense molecules or dsRNA, such as RNAi, antibodies, fragments of antibodies, recombinant antibodies and other such compounds that can serve as
15 drug candidates or lead compounds.

As used herein, a peptidomimetic is a compound that mimics the conformation and certain stereochemical features of the biologically active form of a particular peptide. In general, peptidomimetics are designed to mimic certain desirable properties of a compound, but not the undesirable properties,
20 such as flexibility, that lead to a loss of a biologically active conformation and bond breakdown. Peptidomimetics may be prepared from biologically active compounds by replacing certain groups or bonds that contribute to the undesirable properties with bioisosteres. Bioisosteres are known to those of skill in the art. For example the methylene bioisostere CH_2S has been used as an
25 amide replacement in enkephalin analogs (see, *e.g.*, Spatola (1983) pp. 267-357 in Chemistry and Biochemistry of Amino Acids, Peptides, and Proteins, Weistein, Ed. volume 7, Marcel Dekker, New York). Morphine, which can be administered orally, is a compound that is a peptidomimetic of the peptide endorphin. For purposes herein, cyclic peptides are included among peptidomimetics.

30 As used herein, a promoter region or promoter element refers to a segment of DNA or RNA that controls transcription of the DNA or RNA to which it is operatively linked. The promoter region includes specific sequences that are

-38-

sufficient for RNA polymerase recognition, binding and transcription initiation. This portion of the promoter region is referred to as the promoter. In addition, the promoter region includes sequences that modulate this recognition, binding and transcription initiation activity of RNA polymerase. These sequences can be

5 *cis* acting or can be responsive to *trans* acting factors. Promoters, depending upon the nature of the regulation, can be constitutive or regulated. Exemplary promoters contemplated for use in prokaryotes include the bacteriophage T7 and T3 promoters.

As used herein, a receptor refers to a molecule that has an affinity for a

10 given ligand. Receptors can be naturally-occurring or synthetic molecules. Receptors can also be referred to in the art as anti-ligands. As used herein, the receptor and anti-ligand are interchangeable. Receptors can be used in their unaltered state or as aggregates with other species. Receptors can be attached, covalently or noncovalently, or in physical contact with, to a binding member,

15 either directly or indirectly via a specific binding substance or linker. Examples of receptors, include, but are not limited to: antibodies, cell membrane receptors surface receptors and internalizing receptors, monoclonal antibodies and antisera reactive with specific antigenic determinants [such as on viruses, cells, or other materials], drugs, polynucleotides, nucleic acids, peptides, cofactors, lectins,

20 sugars, polysaccharides, cells, cellular membranes, and organelles.

Examples of receptors and applications using such receptors, include but are not restricted to:

- a) enzymes: specific transport proteins or enzymes essential to survival of microorganisms, which could serve as targets for antibiotic [ligand] selection;
- 25 b) antibodies: identification of a ligand-binding site on the antibody molecule that combines with the epitope of an antigen of interest can be investigated; determination of a sequence that mimics an antigenic epitope can lead to the development of vaccines of which the immunogen is based on one or more of such sequences or lead to the development of related diagnostic agents
- 30 or compounds useful in therapeutic treatments such as for auto-immune diseases
- c) nucleic acids: identification of ligand, such as protein or RNA, binding sites;

d) catalytic polypeptides: polymers, including polypeptides, that are capable of promoting a chemical reaction involving the conversion of one or more reactants to one or more products; such polypeptides generally include a binding site specific for at least one reactant or reaction intermediate and an active functionality proximate to the binding site, in which the functionality is capable of chemically modifying the bound reactant (see, *e.g.*, U.S. Patent No. 5,215,899);

e) hormone receptors: determination of the ligands that bind with high affinity to a receptor is useful in the development of hormone replacement therapies; for example, identification of ligands that bind to such receptors can lead to the development of drugs to control blood pressure; and

f) opiate receptors: determination of ligands that bind to the opiate receptors in the brain is useful in the development of less-addictive replacements for morphine and related drugs.

As used herein, sample refers to anything which can contain an analyte for which an analyte assay is desired. The sample can be a biological sample, such as a biological fluid or a biological tissue. Examples of biological fluids include urine, blood, plasma, serum, saliva, semen, stool, sputum, cerebral spinal fluid, tears, mucus, sperm, amniotic fluid or the like. Biological tissues are aggregate of cells, usually of a particular kind together with their intercellular substance that form one of the structural materials of a human, animal, plant, bacterial, fungal or viral structure, including connective, epithelium, muscle and nerve tissues. Examples of biological tissues also include organs, tumors, lymph nodes, arteries and individual cell(s).

As used herein: stringency of hybridization in determining percentage mismatch is as follows:

- 1) high stringency: 0.1 x SSPE, 0.1% SDS, 65°C
- 2) medium stringency: 0.2 x SSPE, 0.1% SDS, 50°C
- 3) low stringency: 1.0 x SSPE, 0.1% SDS, 50°C

Those of skill in this art know that the washing step selects for stable hybrids and also know the ingredients of SSPE (see, *e.g.*, Sambrook, E.F. Fritsch, T. Maniatis, in: *Molecular Cloning, A Laboratory Manual*, Cold Spring

-40-

Harbor Laboratory Press (1989), vol. 3, p. B.13, see, also, numerous catalogs that describe commonly used laboratory solutions). SSPE is pH 7.4 phosphate-buffered 0.18 NaCl. Further, those of skill in the art recognize that the stability of hybrids is determined by T_m , which is a function of the sodium ion

- 5 concentration and temperature ($T_m = 81.5^\circ \text{C} - 16.6(\log_{10}[\text{Na}^+]) + 0.41(\% \text{G} + \text{C}) - 600/l$), so that the only parameters in the wash conditions critical to hybrid stability are sodium ion concentration in the SSPE (or SSC) and temperature.

- It is understood that equivalent stringencies can be achieved using alternative buffers, salts and temperatures. By way of example and not
10 limitation, procedures using conditions of low stringency are as follows (see also Shilo and Weinberg, *Proc. Natl. Acad. Sci. USA* 78:6789-6792 (1981)): Filters containing DNA are pretreated for 6 hours at 40°C in a solution containing 35% formamide, 5X SSC, 50 mM Tris-HCl (pH 7.5), 5 mM EDTA, 0.1% PVP, 0.1% Ficoll, 1% BSA, and 500 $\mu\text{g/ml}$ denatured salmon sperm DNA (10X SSC is 1.5
15 M sodium chloride, and 0.15 M sodium citrate, adjusted to a pH of 7).

- Hybridizations are carried out in the same solution with the following modifications: 0.02% PVP, 0.02% Ficoll, 0.2% BSA, 100 $\mu\text{g/ml}$ salmon sperm DNA, 10% (wt/vol) dextran sulfate, and $5\text{-}20 \times 10^6$ cpm ^{32}P -labeled probe is used. Filters are incubated in hybridization mixture for 18-20 hours at 40°C ,
20 and then washed for 1.5 hours at 55°C in a solution containing 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 0.1% SDS. The wash solution is replaced with fresh solution and incubated an additional 1.5 hours at 60°C . Filters are blotted dry and exposed for autoradiography. If necessary, filters are washed for a third time at $65\text{-}68^\circ \text{C}$ and reexposed to film. Other conditions of low
25 stringency which can be used are well known in the art (*e.g.*, as employed for cross-species hybridizations).

- By way of example and not way of limitation, procedures using conditions of moderate stringency include, for example, but are not limited to, procedures using such conditions of moderate stringency are as follows: Filters
30 containing DNA are pretreated for 6 hours at 55°C in a solution containing 6X SSC, 5X Denhart's solution, 0.5% SDS and 100 $\mu\text{g/ml}$ denatured salmon sperm DNA. Hybridizations are carried out in the same solution and $5\text{-}20 \times 10^6$ cpm

-41-

³²P-labeled probe is used. Filters are incubated in hybridization mixture for 18-20 hours at 55°C, and then washed twice for 30 minutes at 60°C in a solution containing 1X SSC and 0.1% SDS. Filters are blotted dry and exposed for autoradiography. Other conditions of moderate stringency which can be used
5 are well-known in the art. Washing of filters is done at 37°C for 1 hour in a solution containing 2X SSC, 0.1% SDS.

By way of example and not way of limitation, procedures using conditions of high stringency are as follows: Prehybridization of filters containing DNA is carried out for 8 hours to overnight at 65°C in buffer composed of 6X SSC,
10 50 mM Tris-HCl (pH 7.5), 1 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.02% BSA, and 500 µg/ml denatured salmon sperm DNA. Filters are hybridized for 48 hours at 65°C in prehybridization mixture containing 100 µg/ml denatured salmon sperm DNA and 5-20 X 10⁶ cpm of ³²P-labeled probe. Washing of filters is done at 37°C for 1 hour in a solution containing 2X SSC, 0.01% PVP, 0.01% Ficoll,
15 and 0.01% BSA. This is followed by a wash in 0.1X SSC at 50°C for 45 minutes before autoradiography. Other conditions of high stringency which can be used are well known in the art.

The term substantially identical or substantially homologous or similar varies with the context as understood by those skilled in the relevant art and
20 generally means at least 60% or 70%, preferably means at least 80%, 85% or more preferably at least 90%, and most preferably at least 95% identity.

As used herein, substantially identical to a product means sufficiently similar so that the property of interest is sufficiently unchanged so that the substantially identical product can be used in place of the product.

25 As used herein, substantially pure means sufficiently homogeneous to appear free of readily detectable impurities as determined by standard methods of analysis, such as thin layer chromatography (TLC), gel electrophoresis and high performance liquid chromatography (HPLC), used by those of skill in the art to assess such purity, or sufficiently pure such that further purification would
30 not detectably alter the physical and chemical properties, such as enzymatic and biological activities, of the substance. Methods for purification of the compounds to produce substantially chemically pure compounds are known to

-42-

those of skill in the art. A substantially chemically pure compound can, however, be a mixture of stereoisomers or isomers. In such instances, further purification might increase the specific activity of the compound.

As used herein, target cell refers to a cell that expresses an SP *in vivo*.

5 As used herein, test substance (or test compound) refers to a chemically defined compound (*e.g.*, organic molecules, inorganic molecules, organic/inorganic molecules, proteins, peptides, nucleic acids, oligonucleotides, lipids, polysaccharides, saccharides, or hybrids among these molecules such as glycoproteins, etc.) or mixtures of compounds (*e.g.*, a library of test compounds, 10 natural extracts or culture supernatants, etc.) whose effect on an SP, particularly a single chain form that includes the protease domain or a sufficient portion thereof for activity, as determined by an *in vitro* method, such as the assays provided herein.

As used herein, the terms a therapeutic agent, therapeutic regimen, 15 radioprotectant or chemotherapeutic mean conventional drugs and drug therapies, including vaccines, which are known to those skilled in the art. Radiotherapeutic agents are well known in the art.

As used herein, treatment means any manner in which the symptoms of a condition, disorder or disease are ameliorated or otherwise beneficially altered. 20 Treatment also encompasses any pharmaceutical use of the compositions herein.

As used herein, vector (or plasmid) refers to discrete elements that are used to introduce heterologous nucleic acid into cells for either expression or replication thereof. The vectors typically remain episomal, but can be designed to effect integration of a gene or portion thereof into a chromosome of the 25 genome. Also contemplated are vectors that are artificial chromosomes, such as yeast artificial chromosomes and mammalian artificial chromosomes. Selection and use of such vehicles are well known to those of skill in the art. An expression vector includes vectors capable of expressing DNA that is operatively linked with regulatory sequences, such as promoter regions, that are capable of 30 effecting expression of such DNA fragments. Thus, an expression vector refers to a recombinant DNA or RNA construct, such as a plasmid, a phage, recombinant virus or other vector that, upon introduction into an appropriate

-43-

host cell, results in expression of the cloned DNA. Appropriate expression vectors are well known to those of skill in the art and include those that are replicable in eukaryotic cells and/or prokaryotic cells and those that remain episomal or those which integrate into the host cell genome.

5 As used herein, protein binding sequence refers to a protein or peptide sequence that is capable of specific binding to other protein or peptide sequences generally, to a set of protein or peptide sequences or to a particular protein or peptide sequence.

10 As used herein, epitope tag refers to a short stretch of amino acid residues corresponding to an epitope to facilitate subsequent biochemical and immunological analysis of the epitope tagged protein or peptide. Epitope tagging is achieved by including the sequence of the epitope tag to the protein-encoding sequence in an appropriate expression vector. Epitope tagged proteins can be affinity purified using highly specific antibodies raised against the tags.

15 As used herein, metal binding sequence refers to a protein or peptide sequence that is capable of specific binding to metal ions generally, to a set of metal ions or to a particular metal ion.

 As used herein, a combination refers to any association between two or among more items.

20 As used herein, a composition refers to a any mixture. It can be a solution, a suspension, liquid, powder, a paste, aqueous, non-aqueous or any combination thereof.

 As used herein, fluid refers to any composition that can flow. Fluids thus encompass compositions that are in the form of semi-solids, pastes, solutions, 25 aqueous mixtures, gels, lotions, creams and other such compositions.

 As used herein, a cellular extract refers to a preparation or fraction which is made from a lysed or disrupted cell.

 As used herein, an agent is said to be randomly selected when the agent is chosen randomly without considering the specific sequences involved in the 30 association of a protein alone or with its associated substrates, binding partners, etc. An example of randomly selected agents is the use a chemical library or a

-44-

peptide combinatorial library, or a growth broth of an organism or conditioned medium.

As used herein, an agent is said to be rationally selected or designed when the agent is chosen on a non-random basis which takes into account the sequence of the target site and/or its conformation in connection with the agent's action. As described in the Examples, there are proposed binding sites for serine protease and (catalytic) sites in the protein having SEQ ID NO:3 or SEQ ID NO:4. Agents can be rationally selected or rationally designed by utilizing the peptide sequences that make up these sites. For example, a rationally selected peptide agent can be a peptide whose amino acid sequence is identical to the ATP or calmodulin binding sites or domains.

For clarity of disclosure, and not by way of limitation, the detailed description is divided into the subsections that follow.

B. MTSP9 polypeptides, muteins, derivatives and analogs thereof

15 MTSPs

The MTSPs are a family of transmembrane serine proteases that are found in mammals and also other species. MTSPs are of interest because they appear to be expressed and/or activated at different levels in tumor cells from normal cells, or have functional activity that is different in tumor cells from normal cells, such as by an alteration in a substrate therefor, or a cofactor or a receptor.

The MTSPs share a number of common structural features including: a proteolytic extracellular C-terminal domain; a transmembrane domain, with a hydrophobic domain near the N-terminus; a short cytoplasmic domain; and a variable length stem region that may contain additional modular domains. The proteolytic domains share sequence homology including conserved His, Asp, and Ser residues necessary for catalytic activity that are present in conserved motifs. The MTSPs are normally synthesized as zymogens and can be activated to two-chain forms by cleavage. It is shown herein that the single chain proteolytic domain can function *in vitro* and, hence is useful in *in vitro* assays for identifying agents that modulate the activity of members of this family.

-45-

For purposes herein, the protease domain of the MTSP does not have to result from activation cleavage, which produces a two chain activated product, but rather includes single chain polypeptides where the N-terminii include the consensus sequence ↓VVGG, ↓IVGG, ↓VGLL, ↓ILGG, ↓IVQG or ↓IVNG ↓IASG
5 or other such motif. Such polypeptides, although not the result of activation cleavage and not two-chain forms, exhibit proteolytic (catalytic) activity. These protease domain polypeptides are used in assays to screen for agents that modulate the activity of the MTSP9.

The MTSP family is a target for therapeutic intervention and also some
10 members can serve as diagnostic markers for tumor development, growth and/or progression. As discussed, the members of this family are involved in proteolytic processes that are implicated in tumor development, growth and/or progression. This implication is based upon their functions as proteolytic enzymes in processes related to ECM degradation and/or remodeling and
15 activation of pro-growth factors, pro-hormones or pro-angiogenic compounds. In addition, their levels of expression or level of activation or their apparent activity resulting from substrate levels or alterations in substrates and levels thereof differs in tumor cells and non-tumor cells in the same tissue. Similarly the level of co-factors or receptors for these proteases can vary between tumor and non-
20 tumor cells. Hence, protocols and treatments that alter their activity, such as their proteolytic activities and roles in signal transduction, and/or their expression, such as by contacting them with a compound that modulates their activity and/or expression, could impact tumor development, growth and/or progression. Also, in some instances, the level of activation and/or expression
25 can be altered in tumors, such as lung carcinoma, colon adenocarcinoma and ovarian carcinoma.

MTSP9

MTSP9 is of interest because it is expressed or is active in tumor cells. The MTSP provided herein can serve as a diagnostic marker for particular
30 tumors, by virtue of a level of activity and/or expression or function in a subject (i.e. a mammal, particularly a human) with neoplastic disease, compared to a subject or subjects that do not have the neoplastic disease. In addition,

-46-

detection of activity (and/or expression) in a particular tissue can be indicative of neoplastic disease. It is shown herein, that MTSP9s provided herein are expressed and/or activated in certain tumors; hence their activation or expression can serve as a diagnostic marker for tumor development, growth and/or progression. In other instances, the MTSP polypeptide can exhibit altered activity by virtue of a change in activity or expression of a co-factor, a substrate or a receptor. In addition, in some instances, these MTSPs and/or variants thereof can be shed from cell surfaces. Detection of the shed MTSPs, particularly the extracellular protease domains, in body fluids, such as serum, blood, saliva, cerebral spinal fluid, synovial fluid and interstitial fluids, urine, sweat and other such fluids and secretions, can serve as a diagnostic tumor marker. In particular, detection of higher levels of such shed polypeptides in a subject compared to a subject known not to have any neoplastic disease or compared to earlier samples from the same subject, can be indicative of neoplastic disease in the subject.

Polypeptides and muteins

Provided herein are isolated substantially pure single chain and two chain polypeptides that contain the protease domain of an MTSP9. The polypeptides also can include other non-MTSP sequences of amino acids, but includes the protease domain or a sufficient portion thereof to exhibit catalytic activity in any *in vitro* assay that assess such protease activity, such as any provided herein.

MTSP9 polypeptides provided herein are expressed or activated by or in tumor cells, typically at a level that differs from the level in which they are expressed by or activated in a non-tumor cell of the same type. Hence, for example, if the MTSP is expressed in an cervical tumor cell, it is expressed or active at a different level from in non-tumor cervical cells. MTSP9 expression or activation can be indicative of cervical, lung, esophageal, colon, prostate, uterine, pancreatic, breast and other tumors.

Isolated, substantially pure proteases that include protease domains or a catalytically active portion thereof are provided. Provided are single chain forms and two chain forms of the MTSP9. The protease domains can be included in a longer protein, and such longer protein is optionally the MTSP9 zymogen.

-47-

Exemplary MTSP9-encoding nucleic acid and protein sequences of a protease domain are set forth in SEQ ID Nos. 5 and 6, and a full-length protein and encoding nucleic acid sequence is set forth in SEQ ID Nos. 18 and 17. Thus, an MTSP9 polypeptide includes the sequence of amino acids set forth in SEQ ID Nos. 6, 16 or 18. Smaller portions thereof that retain protease activity are contemplated. The protease domain thereof is set forth in SEQ ID No. 16.

Substantially purified MTSP9 protease is encoded by a nucleic acid that hybridizes to a nucleic acid molecule containing the protease domain encoded by the nucleotide sequence set forth in any of SEQ. ID Nos. 5 and 17 under at least moderate, generally high, stringency conditions, such that the protease domain encoding nucleic acid thereof hybridizes along its full-length or at least 70%, 80% or 90% of the full-length. In certain embodiments the substantially purified MTSP protease is a single chain polypeptide that includes substantially the sequence of amino acids set forth in SEQ ID No. 6, 18 or the protease domain portion thereof, or a catalytically active portion thereof.

Also included are substantially purified MTSP9 zymogens, activated two chain forms, single chain protease domains and two chain protease domains. These polypeptides are encoded by a nucleic acid that includes sequence encoding a protease domain that exhibits proteolytic activity and that hybridizes to a nucleic acid molecule having a nucleotide sequence set forth in SEQ ID No. 5 or 7, typically under moderate, generally under high stringency, conditions and generally along the full-length or along at least about 70%, 80% or 90% of the full-length (or substantially the full-length) of the protease domain. Splice variants are also contemplated herein.

Protease domains

MTSP protease domains include the single chain protease domains of MTSP9. Provided are the protease domains or proteins that include a portion of an MTSP that is the protease domain of any MTSP, particularly a MTSP9. The protein can also include other non-MTSP sequences of amino acids, but includes the protease domain or a sufficient portion thereof to exhibit catalytic activity in any *in vitro* assay that assess such protease activity, such as any provided

-48-

herein. Also provided are two chain activated forms of the full length protease and also two chain forms of the protease domain.

Thus, isolated, substantially pure proteases that include the protease domains or catalytically active portions thereof as single chain forms of SPs are provided. The protease domains can be included in a longer protein, and such longer protein is optionally the MTSP9 zymogen

In particular, exemplary protease domains include at least a sufficient portion of sequences of amino acids set forth as amino acids 206-438 in SEQ ID No. 16 (encoded by nucleotides in SEQ ID No. 15 and 17).

As noted, the protease domains of an MTSP are single-chain polypeptides or two-chain polypeptides, with an N-terminus (such as IV, VV, IL and II) generated at the cleavage site (generally having the consensus sequence R↓VVG, R↓IVG, R↓IVQ, R↓IVNG, R↓ILG, R↓VGL, R↓ILG or a variation thereof; an N-terminus R↓V or R↓I, where the arrow represents the cleavage point) when the zymogen is activated. The protease domain of an exemplary MTSP9, produced is produced by activation cleavage between R₁₈₅ and I₁₈₆ (R↓I) includes the sequence R↓IASG, as set forth in SEQ ID Nos. 17 and 18.

Muteins and derivatives

Full-length MTSP9, zymogen and activated forms thereof and MTSP9 protease domains, portions thereof, and muteins and derivatives of such polypeptides are provided. The domains, fragments, derivatives or analogs of an MTSP9 that are functionally active are capable of exhibiting one or more functional activities associated with the MTSP9 polypeptide, such as serine protease activity, immunogenicity and antigenicity, are provided.

Among the derivatives are those based on animal MTSP9s, including, but are not limited to, rodent, such as mouse and rat; fowl, such as chicken; ruminants, such as goats, cows, deer, sheep; ovine, such as pigs; and humans. For example, MTSP9 derivatives can be made by altering their sequences by substitutions, additions or deletions. MTSP9 derivatives include, but are not limited to, those containing, as a primary amino acid sequence, all or part of the amino acid sequence of MTSP9, including altered sequences in which functionally equivalent amino acid residues are substituted for residues within

the sequence resulting in a silent change. For example, one or more amino acid residues within the sequence can be substituted by another amino acid of a similar polarity which acts as a functional equivalent, resulting in a silent alteration. Substitutes for an amino acid within the sequence can be selected
5 from other members of the class to which the amino acid belongs. For example, the nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan and methionine. The polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine. The positively charged (basic) amino acids include arginine,
10 lysine and histidine. The negatively charged (acidic) amino acids include aspartic acid and glutamic acid (see, *e.g.*, Table 1). Muteins of the MTSP9 or a domain thereof, such as a protease domain, in which up to about 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 85%, 90% or 95% of the amino acids are replaced with another amino acid are provided. Generally such muteins retain at
15 least about 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% or 90% of the protease activity the unmutated protein.

Muteins of the MTSP9 or a domain thereof, such as a protease domain, in which up to about 10%, 20%, 30%, 40%, 50%, 60%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%,
20 95%, 96%, 97%, 98% or 99% of the amino acids are replaced with another amino acid are provided. Generally such muteins retain at least about 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% or 90% of the protease activity the unmutated protein.

Included among the polypeptides provided herein are the MTSP9 protease
25 domain or a polypeptide with amino acid changes such that the specificity and protease activity remains substantially unchanged or changed (increased or decreased) by a specified percentage, such as 10, 20, 30, 40, 50%. In particular, a substantially purified mammalian MTSP polypeptide is provided that has a transmembrane domain and can additionally include a transmembrane (TM)
30 domain, a SEA domain and a serine protease catalytic domain is provided.

Also provided is a substantially purified protein containing a sequence of amino acids that has at least 60%, 70%, 75%, 80%, 81%, 82%, 83%, 84%,

-50-

85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identity to the MTSP9 where the percentage identity is determined using standard algorithms and gap penalties that maximize the percentage identity. The human MTSP9 polypeptide is included, although other mammalian
5 MTSP9 polypeptides are contemplated. The precise percentage of identity can be specified if needed.

Muteins in which one or more of the Cys residues, particularly, a residue that is paired in the activated two form, but unpaired in the protease domain alone is/are replaced with any amino acid, typically, although not necessarily, a
10 conservative amino acid residue, such as Ser, are contemplated. Disulfide bonds pairing in MTSP9 is as follows: C₁₇₅-C₂₉₂, C₂₁₂-C₂₂₈, C₃₃₇-C₃₅₃, C₃₆₄-C₃₉₃. The Cys₂₉₂ is in the protease domain and is unpaired in the single chain form of the protease domain. Muteins of MTSP9, particularly those in which Cys residues, such as the Cys₂₉₂ in the single chain protease domain, is replaced with another
15 amino acid, such as Ser, Gly or Ala, that does not eliminate the activity, are provided. Also provided are substantially purified MTSP9 polypeptides and functional domains thereof, including catalytically active domains and portions, that have at least about 60%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%
20 or 99% sequence identity with a protease domain that includes the sequence of amino acids set forth in SEQ ID No. 16 or a catalytically active portion thereof or with a protease that includes the sequence of amino acids set forth in SEQ ID No. 18 and domains thereof.

Muteins of the protein are also provided in which amino acids are
25 replaced with other amino acids. Among the muteins are those in which the Cys residues, is/are replaced typically with a conservative amino acid residues, such as a serine. Such muteins are also provided herein. Muteins in which 10%, 20%, 30%, 35%, 40%, 45%, 50% or more of the amino acids are replaced but the resulting polypeptide retains at least about 10%, 20%, 30%, 35%, 40%,
30 45%, 50%, 60%, 70%, 80%, 90% or 95% of the catalytic activity as the unmodified form for the same substrate.

-51-

Mutagens can be made by making conservative amino acid substitutions and also non-conservative amino acid substitutions. For example, amino acid substitutions that desirably alter properties of the proteins can be made. In one embodiment, mutations that prevent degradation of the polypeptide can be made. Many proteases cleave after basic residues, such as R and K; to eliminate such cleavage, the basic residue is replaced with a non-basic residue. Also, non-conservative changes at amino acids outside of the protease domain can be effected without altering protease activity. Non-conservative changes at amino acids that are responsible for activities other than protease activity may be desirable. For example, interaction of the protease with an inhibitor can be blocked while retaining catalytic activity by effecting a non-conservative change at the site of interaction of the inhibitor with the protease. Similarly, receptor binding can be altered without altering catalytic activity by effecting a non-conservative or conservative change at a site of interaction of the receptor with the protease.

Antigenic epitopes that contain at least 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 20, 25, 30, 40, 50, and typically 10-15 amino acids of the MTSP9 polypeptide are provided. These antigenic epitopes are used, for example, to raise antibodies. Antibodies specific for each epitope or combinations thereof and for single and two-chain forms are also provided.

Nucleic acid molecules, vectors and plasmids, cells and expression of MTSP9 polypeptides

Nucleic acid molecules

Due to the degeneracy of nucleotide coding sequences, other nucleic acid sequences which encode substantially the same amino acid sequence as a MTSP9 are contemplated. These include but are not limited to nucleic acid molecules that include all or portions of MTSP9-encoding genes that are altered by the substitution of different codons that encode the amino acid residue within the sequence, thus producing a silent change.

Nucleic acids

Also provided herein are nucleic acid molecules that encode MTSP9 polypeptides and the encoded proteins. In particular, nucleic acid molecules

-52-

encoding MTSP9 from animals, including splice variants thereof are provided. The encoded proteins are also provided. Also provided are functional domains thereof. \ For each of the nucleic acid molecules provided, the nucleic acid can be DNA or RNA or PNA or other nucleic acid analogs or can include non-
5 natural nucleotide bases. Also provided are isolated nucleic acid molecules that include a sequence of nucleotides complementary to the nucleotide sequence encoding an MTSP.

Also provided are nucleic acid molecules that encode a single chain or two chain MTSP proteases that have proteolytic activity in an *in vitro* proteolysis
10 assay and that have at least 60%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity with the full-length of a protease domain of an MTSP9 polypeptide, or that hybridize along their full-length or along at least about 70%, 80% or 90% of the full-length nucleic acid to a nucleic acids that
15 encode a protease domain, particularly under conditions of moderate, generally high, stringency. As above, the encoded polypeptides contain the protease as a single chain; activated forms thereof can be produced and are provided.

In one embodiment, a nucleic acid molecule that encodes an MTSP, designated MTSP9 is provided. The nucleic acid molecule includes the open
20 reading frame in the sequence of nucleotides set forth in SEQ ID No. 17. Also provided are nucleic acid molecules that hybridize under conditions of at least low stringency, moderate stringency, and generally high stringency to the following sequence of nucleic acids (SEQ ID No. 5 or 17) particularly to the open reading frame encompassed by nucleotides that encode a single protease domain
25 thereof, or any domain of MTSP9.

In certain embodiments, the isolated nucleic acid fragment hybridizes to the nucleic acid having the nucleotide sequence set forth in SEQ ID No. 5 or 17 under high stringency conditions, and generally contains the sequence of nucleotides set forth in SEQ ID Nos. 5 or 17. The protein contains a
30 transmembrane domain (TM) and a serine protease domain.

Also provided, are muteins of the nucleic acid molecules that encode polypeptides in which amino acids are replaced with other amino acids. Among

the muteins are those in which the Cys residue-encoding codons, is/are replaced with other amino acid residues, such as a codon encoding a serine. Such muteins are also provided herein. Each of such domains is provided herein as are nucleic acid molecules that include sequences of nucleotides encoding such domains. Some MTSPs can additionally include a LDLR domain, a scavenger-receptor cysteine rich (SRCR) domain and other domains.

The isolated nucleic acid fragment is DNA, including genomic or cDNA, or is RNA, or can include other components, such as protein nucleic acid and other nucleotide analogs. The isolated nucleic acid can include additional components, such as heterologous or native promoters, and other transcriptional and translational regulatory sequences, these genes can be linked to other genes, such as reporter genes or other indicator genes or genes that encode indicators.

Also provided are nucleic acid molecules that hybridize to the above-noted sequences of nucleotides encoding MTSP9 at least at low stringency, moderate stringency, and typically at high stringency, and that encode the protease domain and/or the full-length protein or at least 70%, 80% or 90% of the full-length protease domain or other domains of an MTSP9 or a splice variant or allelic variant thereof. Generally the molecules hybridize under such conditions along their full-length or along at least 70%, 80% or 90% of the full-length for at least one domain and encode at least one domain, such as the protease or extracellular domain, of the polypeptide. In particular, such nucleic acid molecules include any isolated nucleic fragment that encodes at least one domain of a membrane serine protease, that (1) contains a sequence of nucleotides that encodes the protease or a domain thereof, and (2) is selected from among:

- (a) a sequence of nucleotides that encodes the protease or a domain thereof that includes a sequence of nucleotides set forth in SEQ ID Nos. 15 or 17;
- (b) a sequence of nucleotides that encodes such portion or the full-length protease and hybridizes under conditions of high stringency, generally to nucleic acid that is complementary to a mRNA

-54-

transcript present in a mammalian cell that encodes such protein or fragment thereof;

- 5 (a) a sequence of nucleotides that encodes the protease or a domain thereof that includes a sequence of nucleotides set having at least about 60%, 70%, 80%, 90% or 95% sequence identity the the sequence set forth in SEQ ID Nos. 5, 15 or 17;
- (c) a sequence of nucleotides that encodes a transmembrane protease or domain thereof that includes a sequence of amino acids encoded by such portion or the full-length open reading frame;
- 10 (d) a sequence of nucleotides that encodes the protease or a domain thereof that includes a sequence of nucleotides set having at least about 60%, 70%, 80%, 90% or 95% sequence identity the the sequence set forth in SEQ ID Nos. 5, 15 or 17; and
- 15 (e) a sequence of nucleotides that encodes the transmembrane protease that includes a sequence of amino acids encoded by a sequence of nucleotides that encodes such subunit and hybridizes under conditions of low, moderate or high stringency to DNA that is complementary to the mRNA transcript.

The isolated nucleic acids can contain least 10 nucleotides, 25
 20- nucleotides, 50 nucleotides, 100 nucleotides, 150 nucleotides, or 200 nucleotides or more contiguous nucleotides of an MTSP9-encoding sequence, or a full-length SP coding sequence. In another embodiment, the nucleic acids are smaller than 35, 200 or 500 nucleotides in length. Nucleic acids that hybridize to or are complementary to an MTSP9-encoding nucleic acid molecule can be
 25 single or double-stranded. For example, nucleic acids are provided that include a sequence complementary to (specifically are the inverse complement of) at least 10, 25, 50, 100, or 200 nucleotides or the entire coding region of an MTSP9 encoding nucleic acid, particularly the protease domain thereof. For MTSP9 the full-length protein or a domain or active fragment thereof is also provided.

30 **Probes, primers, antisense oligonucleotides and dsRNA**

Also provided are fragments thereof that can be used as probes or primers and that contain at least about 10 nucleotides, 14 nucleotides, generally

-55-

at least about 16 nucleotides, often at least about 30 nucleotides. The length of the probe or primer is a function of the size of the genome probed; the larger the genome, the longer the probe or primer required for specific hybridization to a single site. Those of skill in the art can select appropriately sized probes and
5 primers. Generally probes and primers as described are single-stranded. Double stranded probes and primers can be used, if they are denatured when used.

Probes and primers derived from the nucleic acid molecules are provided. Such probes and primers contain at least 8, 14, 16, 30, 100 or more contiguous nucleotides with identity to contiguous nucleotides of an MTSP9, generally,
10 except for nucleotides 634-751 of SEQ ID No. 5 or nucleotides 1162-1279 of SEQ ID No. 17, and probes of at least 30, 50 or 100 contiguous sequence of nucleotides of SEQ ID No. 5, except for nucleotides 634-734 of SEQ ID No. 5 (except for nucleotides 1162-1262 of SEQ ID No. 18). The probes and primers are optionally labelled with a detectable label, such as a radiolabel or a
15 fluorescent tag, or can be mass differentiated for detection by mass spectrometry or other means.

Also provided is an isolated nucleic acid molecule that includes the sequence of molecules that is complementary to the nucleotide sequence encoding MTSP9 or the portion thereof. Double-stranded RNA (dsRNA), such
20 as RNAi is also provided.

Plasmids, Vectors and Cells

Plasmids and vectors containing the nucleic acid molecules are also provided. Cells containing the vectors, including cells that express the encoded proteins are provided. The cell can be a bacterial cell, a yeast cell, a fungal cell,
25 a plant cell, an insect cell or an animal cell. Methods for producing an MTSP or single chain form of the protease domain thereof by, for example, growing the cell under conditions whereby the encoded MTSP is expressed by the cell, and recovering the expressed protein, are provided herein. As noted, for MTSP9, the full-length zymogens and activated proteins and activated (two chain) protease

-56-

and single chain protease domains are provided. As described herein, the cells are used for expression of the protein, which can be secreted or expressed in the cytoplasm.

As discussed below, the MTSP9 polypeptide, and catalytically active portions thereof, can be expressed on the surface of a cell. In addition, all or portions thereof can be expressed as a secreted protein using the native signal sequence or a heterologous signal. Alternatively, all or portions of the polypeptide can be expressed as inclusion bodies in the cytoplasm and isolated therefrom. The resulting protein can be treated to refold if necessary.

The above discussion provides an overview and some details of the exemplified MTSP9s.

C. Tumor specificity and tissue expression profiles

Each MTSP has a characteristic tissue expression profile; the MTSPs in particular, although not exclusively expressed or activated in tumors, exhibit characteristic tumor tissue expression or activation profiles. In some instances, MTSPs can have different activity in a tumor cell from a non-tumor cell by virtue of a change in a substrate or cofactor or receptor therefor or other factor that would alter the apparent functional activity of the MTSP. Hence each can serve as a diagnostic marker for particular tumors, by virtue of a level of activity and/or expression or function in a subject (i.e. a mammal, particularly a human) with neoplastic disease, compared to a subject or subjects that do not have the neoplastic disease. In addition, detection of activity (and/or expression) in a particular tissue can be indicative of neoplastic disease. Shed MTSPs in body fluids can be indicative of neoplastic disease. Also, by virtue of the activity and/or expression profiles of each, they can serve as therapeutic targets, such as by administration of modulators of the activity thereof, or, as by administration of a prodrug specifically activated by one of the MTSPs.

Tissue expression profiles

MTSP9

The MTSP9 is highly expressed in the esophagus and expressed at a low level in many other tissues. The MTSP9 transcript is found in kidney (adult and fetal), spleen (adult and fetal), placenta, liver (adult and fetal), thymus, peripheral

-57-

blood leukocyte, lung (adult and fetal), pancreas, lymph node, bone marrow, trachea, uterus, prostate, testes, ovary and the gland organs (mammary, adrenal, thyroid, pituitary and salivary). MTSP9 is also expressed in esophagus tumor tissues, in a lung carcinoma and, at a lower level, in a colorectal carcinoma,

5 lymphoma, a cervical carcinoma (HeLaS3) and leukemia cell lines.

D. Identification and isolation of MTSP9 polypeptide genes

The MTSP polypeptides and/or domains thereof, can be obtained by methods well known in the art for protein purification and recombinant protein expression. Any method known to those of skill in the art for identification of
10 nucleic acids that encode desired genes can be used. Any method available in the art can be used to obtain a full-length (*i.e.*, encompassing the entire coding region) cDNA or genomic DNA clone encoding an MTSP polypeptide. For example, the polymerase chain reaction (PCR) can be used to amplify a sequence that is expressed in normal and tumor cells or tissues, *e.g.*, nucleic
15 acids encoding an MTSP9 polypeptide (SEQ. Nos: 5 and 17), in a genomic or cDNA library. Oligonucleotide primers that hybridize to sequences at the 3' and 5' termini of the identified sequences can be used as primers to amplify by PCR sequences from a nucleic acid sample (RNA or DNA), generally a cDNA library, from an appropriate source (*e.g.*, tumor or cancer tissue).

20 PCR can be carried out, *e.g.*, by use of a Perkin-Elmer Cetus thermal cycler and Taq polymerase (Gene Amp™). The DNA being amplified can include mRNA or cDNA or genomic DNA from any eukaryotic species. One can choose to synthesize several different degenerate primers, for use in the PCR reactions. It is also possible to vary the stringency of hybridization conditions used in
25 priming the PCR reactions, to amplify nucleic acid homologs (*e.g.*, to obtain *MTSP polypeptide* sequences from species other than humans or to obtain human sequences with homology to MTSP9 polypeptide) by allowing for greater or lesser degrees of nucleotide sequence similarity between the known nucleotide sequence and the nucleic acid homolog being isolated. For cross-
30 species hybridization, low stringency to moderate stringency conditions are used. For same species hybridization, moderately stringent to highly stringent conditions are used. The conditions can be empirically determined.

-58-

After successful amplification of the nucleic acid containing all or a portion of the identified MTSP polypeptide sequence or of a nucleic acid encoding all or a portion of an MTSP polypeptide homolog, that segment can be molecularly cloned and sequenced, and used as a probe to isolate a complete cDNA or genomic clone. This, in turn, permits the determination of the gene's complete nucleotide sequence, the analysis of its expression, and the production of its protein product for functional analysis. Once the nucleotide sequence is determined, an open reading frame encoding the MTSP polypeptide gene protein product can be determined by any method well known in the art for determining open reading frames, for example, using publicly available computer programs for nucleotide sequence analysis. Once an open reading frame is defined, it is routine to determine the amino acid sequence of the protein encoded by the open reading frame. In this way, the nucleotide sequences of the entire MTSP polypeptide genes as well as the amino acid sequences of MTSP polypeptide proteins and analogs can be identified.

Any eukaryotic cell potentially can serve as the nucleic acid source for the molecular cloning of the MTSP polypeptide gene. The nucleic acids can be isolated from vertebrate, mammalian, human, porcine, bovine, feline, avian, equine, canine, as well as additional primate sources, insects, plants and other organisms. The DNA can be obtained by standard procedures known in the art from cloned DNA (*e.g.*, a DNA "library"), by chemical synthesis, by cDNA cloning, or by the cloning of genomic DNA, or fragments thereof, purified from the desired cell (see, *e.g.*, Sambrook *et al.* 1989, *Molecular Cloning, A Laboratory Manual*, 2d Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York; Glover, D.M. (ed.), 1985, *DNA Cloning: A Practical Approach*, MRL Press, Ltd., Oxford, U.K. Vol. I, II). Clones derived from genomic DNA can contain regulatory and intron DNA regions in addition to coding regions; clones derived from cDNA will contain only exon sequences. For any source, the gene is cloned into a suitable vector for propagation thereof.

In the molecular cloning of the gene from genomic DNA, DNA fragments are generated, some of which will encode the desired gene.

The DNA can be cleaved at specific sites using various restriction enzymes. Alternatively, one can use DNase in the presence of manganese to fragment the DNA, or the DNA can be physically sheared, for example, by sonication. The linear DNA fragments then can be separated according to size by standard
5 techniques, including but not limited to, agarose and polyacrylamide gel electrophoresis and column chromatography.

Once the DNA fragments are generated, identification of the specific DNA fragment containing the desired gene can be accomplished in a number of ways. For example, a portion of the MTSP polypeptide (of any species) gene (*e.g.*, a
10 PCR amplification product obtained as described above or an oligonucleotide having a sequence of a portion of the known nucleotide sequence) or its specific RNA, or a fragment thereof be purified and labeled, and the generated DNA fragments can be screened by nucleic acid hybridization to the labeled probe (Benton and Davis, *Science* 196:180 (1977); Grunstein and Hogness, *Proc. Natl.*
15 *Acad. Sci. U.S.A.* 72:3961 (1975)). Those DNA fragments with substantial homology to the probe will hybridize. It is also possible to identify the appropriate fragment by restriction enzyme digestion(s) and comparison of fragment sizes with those expected according to a known restriction map if such is available or by DNA sequence analysis and comparison to the known
20 nucleotide sequence of MTSP polypeptide. Further selection can be carried out on the basis of the properties of the gene. Alternatively, the presence of the gene can be detected by assays based on the physical, chemical, or immunological properties of its expressed product. For example, cDNA clones, or DNA clones which hybrid-select the proper mRNA, can be selected which
25 produce a protein that, *e.g.*, has similar or identical electrophoretic migration, isoelectric focusing behavior, proteolytic digestion maps, antigenic properties, serine protease activity. If an anti-MTSP polypeptide antibody is available, the protein can be identified by binding of labeled antibody to the putatively MTSP polypeptide synthesizing clones, in an ELISA (enzyme-linked immunosorbent
30 assay)-type procedure.

Alternatives to isolating the MTSP9 polypeptide genomic DNA include, but are not limited to, chemically synthesizing the gene sequence from a known

-60-

sequence or making cDNA to the mRNA that encodes the MTSP polypeptide. For example, RNA for cDNA cloning of the MTSP polypeptide gene can be isolated from cells expressing the protein. The identified and isolated nucleic acids then can be inserted into an appropriate cloning vector. A large number of vector-host systems known in the art can be used. Possible vectors include, but are not limited to, plasmids or modified viruses, but the vector system must be compatible with the host cell used. Such vectors include, but are not limited to, bacteriophages such as lambda derivatives, or plasmids such as pBR322 or pUC plasmid derivatives or the Bluescript vector (Stratagene, La Jolla, CA). The insertion into a cloning vector can, for example, be accomplished by ligating the DNA fragment into a cloning vector which has complementary cohesive termini. If the complementary restriction sites used to fragment the DNA are not present in the cloning vector, the ends of the DNA molecules can be enzymatically modified. Alternatively, any site desired can be produced by ligating nucleotide sequences (linkers) onto the DNA termini; these ligated linkers can include specific chemically synthesized oligonucleotides encoding restriction endonuclease recognition sequences. In an alternative method, the cleaved vector and MTSP polypeptide gene can be modified by homopolymeric tailing. Recombinant molecules can be introduced into host cells via transformation, transfection, infection, electroporation, calcium precipitation and other methods, so that many copies of the gene sequence are generated.

In specific embodiments, transformation of host cells with recombinant DNA molecules that incorporate the isolated MTSP polypeptide gene, cDNA, or synthesized DNA sequence enables generation of multiple copies of the gene. Thus, the gene can be obtained in large quantities by growing transformants, isolating the recombinant DNA molecules from the transformants and, when necessary, retrieving the inserted gene from the isolated recombinant DNA.

E. Vectors, plasmids and cells that contain nucleic acids encoding an MTSP polypeptide or protease domain thereof and expression of MTSP polypeptides

Vectors and cells

For recombinant expression of one or more of the MTSP polypeptides, the nucleic acid containing all or a portion of the nucleotide sequence encoding the

-61-

MTSP polypeptide can be inserted into an appropriate expression vector, *i.e.*, a vector that contains the necessary elements for the transcription and translation of the inserted protein coding sequence. The necessary transcriptional and translational signals can also be supplied by the native promoter for MTSP

5 genes, and/or their flanking regions.

Also provided are vectors that contain nucleic acid encoding the MTSPs. Cells containing the vectors are also provided. The cells include eukaryotic and prokaryotic cells, and the vectors are any suitable for use therein.

Prokaryotic and eukaryotic cells, including endothelial cells, containing the
10 vectors are provided. Such cells include bacterial cells, yeast cells, fungal cells, plant cells, insect cells and animal cells. The cells are used to produce an MTSP polypeptide or protease domain thereof by (a) growing the above-described cells under conditions whereby the encoded MTSP polypeptide or protease domain of the MTSP polypeptide is expressed by the cell, and then (b)
15 recovering the expressed protease domain protein. In the exemplified embodiments, the protease domain is secreted into the medium.

In one embodiment, the vectors include a sequence of nucleotides that encodes a polypeptide that has protease activity and contains all or a portion of only the protease domain, or multiple copies thereof, of an SP protein are
20 provided. Also provided are vectors that comprise a sequence of nucleotides that encodes the protease domain and additional portions of an SP protein up to and including a full length SP protein, as well as multiple copies thereof, are also provided. The vectors can be selected for expression of the SP protein or protease domain thereof in the cell or such that the SP protein is expressed as a secreted
25 protein. Alternatively, the vectors can include signals necessary for secretion of encoded proteins. When the protease domain is expressed the nucleic acid is linked to nucleic acid encoding a secretion signal, such as the *Saccharomyces cerevisiae* α mating factor signal sequence or a portion thereof, or the native signal sequence.

30 A variety of host-vector systems can be used to express the protein coding sequence. These include but are not limited to mammalian cell systems infected with virus (*e.g.* vaccinia virus, adenovirus, etc.); insect cell systems

-62-

infected with virus (*e.g.* baculovirus); microorganisms such as yeast containing yeast vectors; or bacteria transformed with bacteriophage, DNA, plasmid DNA, or cosmid DNA. The expression elements of vectors vary in their strengths and specificities. Depending on the host-vector system used, any one of a number
5 of suitable transcription and translation elements can be used.

Any methods known to those of skill in the art for the insertion of nucleic acid fragments into a vector can be used to construct expression vectors containing a chimeric gene containing appropriate transcriptional/translational control signals and protein coding sequences. These methods can include *in*
10 *vitro* recombinant DNA and synthetic techniques and *in vivo* recombinants (genetic recombination). Expression of nucleic acid sequences encoding MTSP polypeptide, or domains, derivatives, fragments or homologs thereof, can be regulated by a second nucleic acid sequence so that the genes or fragments thereof are expressed in a host transformed with the recombinant DNA
15 molecule(s). For example, expression of the proteins can be controlled by any promoter/enhancer known in the art. In a specific embodiment, the promoter is not native to the genes for MTSP polypeptide. Promoters which can be used include but are not limited to the SV40 early promoter (Bernoist and Chambon, *Nature* 290:304-310 (1981)), the promoter contained in the 3' long terminal
20 repeat of Rous sarcoma virus (Yamamoto et al., *Cell* 22:787-797 (1980)), the herpes thymidine kinase promoter (Wagner et al., *Proc. Natl. Acad. Sci. USA* 78:1441-1445 (1981)), the regulatory sequences of the metallothionein gene (Brinster et al., *Nature* 296:39-42 (1982)); prokaryotic expression vectors such as the β -lactamase promoter (Villa-Kamaroff et al., *Proc. Natl. Acad. Sci. USA*
25 75:3727-3731 (1978)) or the *tac* promoter (DeBoer et al., *Proc. Natl. Acad. Sci. USA* 80:21-25 (1983)); see also "Useful Proteins from Recombinant Bacteria": in *Scientific American* 242:79-94 (1980)); plant expression vectors containing the nopaline synthetase promoter (Herrar-Estrella et al., *Nature* 303:209-213 (1984)) or the cauliflower mosaic virus 35S RNA promoter (Gardner et al., *Nucleic Acids*
30 *Res.* 9:2871 (1981)), and the promoter of the photosynthetic enzyme ribulose biphosphate carboxylase (Herrera-Estrella et al., *Nature* 310:115-120 (1984)); promoter elements from yeast and other fungi such as the Gal4 promoter, the

alcohol dehydrogenase promoter, the phosphoglycerol kinase promoter, the alkaline phosphatase promoter, and the following animal transcriptional control regions that exhibit tissue specificity and have been used in transgenic animals: elastase I gene control region which is active in pancreatic acinar cells (Swift et al., *Cell* 38:639-646 (1984); Ornitz et al., *Cold Spring Harbor Symp. Quant. Biol.* 50:399-409 (1986); MacDonald, *Hepatology* 7:425-515 (1987)); insulin gene control region which is active in pancreatic beta cells (Hanahan et al., *Nature* 315:115-122 (1985)), immunoglobulin gene control region which is active in lymphoid cells (Grosschedl et al., *Cell* 38:647-658 (1984); Adams et al., *Nature* 318:533-538 (1985); Alexander et al., *Mol. Cell Biol.* 7:1436-1444 (1987)), mouse mammary tumor virus control region which is active in testicular, breast, lymphoid and mast cells (Leder et al., *Cell* 45:485-495 (1986)), albumin gene control region which is active in liver (Pinckert et al., *Genes and Devel.* 1:268-276 (1987)), alpha-fetoprotein gene control region which is active in liver (Krumlauf et al., *Mol. Cell Biol.* 5:1639-1648 (1985); Hammer et al., *Science* 235:53-58 (1987)), alpha-1 antitrypsin gene control region which is active in liver (Kelsey et al., *Genes and Devel.* 1:161-171 (1987)), beta globin gene control region which is active in myeloid cells (Mogam et al., *Nature* 315:338-340 (1985); Kollias et al., *Cell* 46:89-94 (1986)), myelin basic protein gene control region which is active in oligodendrocyte cells of the brain (Readhead et al., *Cell* 48:703-712 (1987)), myosin light chain-2 gene control region which is active in skeletal muscle (Sani, *Nature* 314:283-286 (1985)), and gonadotrophic releasing hormone gene control region which is active in gonadotrophs of the hypothalamus (Mason et al., *Science* 234:1372-1378 (1986)).

In a specific embodiment, a vector is used that contains a promoter operably linked to nucleic acids encoding an MTSP polypeptide, or a domain, fragment, derivative or homolog, thereof, one or more origins of replication, and optionally, one or more selectable markers (*e.g.*, an antibiotic resistance gene). Expression vectors containing the coding sequences, or portions thereof, of an MTSP polypeptide, is made, for example, by subcloning the coding portions into the EcoRI restriction site of each of the three pGEX vectors (glutathione S-transferase expression vectors (Smith and Johnson, *Gene* 7:31-40 (1988)). This

-64-

allows for the expression of products in the correct reading frame. Exemplary vectors and systems for expression of the protease domains of the MTSP polypeptides include the well-known *Pichia* vectors (available, for example, from Invitrogen, San Diego, CA), particularly those designed for secretion of the
5 encoded proteins. The protein can also be expressed cytoplasmically, such as in the inclusion bodies. One exemplary vector is described in the EXAMPLES.

Plasmids for transformation of *E. coli* cells, include, for example, the pET expression vectors (see, U.S patent 4,952,496; available from NOVAGEN, Madison, WI; see, also literature published by Novagen describing the system).
10 Such plasmids include pET 11a, which contains the T7lac promoter, T7 terminator, the inducible *E. coli* lac operator, and the lac repressor gene; pET 12a-c, which contains the T7 promoter, T7 terminator, and the *E. coli* ompT secretion signal; and pET 15b and pET19b (NOVAGEN, Madison, WI), which contain a His-Tag™ leader sequence for use in purification with a His column and
15 a thrombin cleavage site that permits cleavage following purification over the column; the T7-lac promoter region and the T7 terminator.

The vectors are introduced into host cells, such as *Pichia* cells and bacterial cells, such as *E. coli*, and the proteins expressed therein. Exemplary *Pichia* strains, include, for example, GS115. Exemplary bacterial hosts contain
20 chromosomal copies of DNA encoding T7 RNA polymerase operably linked to an inducible promoter, such as the lacUV promoter (see, U.S. Patent No. 4,952,496). Such hosts include, but are not limited to, the lysogenic *E. coli* strain BL21(DE3).

Expression and production of proteins

25 The MTSP domains, derivatives and analogs can be produced by various methods known in the art. For example, once a recombinant cell expressing an MTSP polypeptide, or a domain, fragment or derivative thereof, is identified, the individual gene product can be isolated and analyzed. This is achieved by assays based on the physical and/or functional properties of the protein,
30 including, but not limited to, radioactive labeling of the product followed by analysis by gel electrophoresis, immunoassay, cross-linking to marker-labeled product, and assays of proteolytic activity.

-65-

The MTSP polypeptides can be isolated and purified by standard methods known in the art (either from natural sources or recombinant host cells expressing the complexes or proteins), including but not restricted to column chromatography (*e.g.*, ion exchange, affinity, gel exclusion, reversed-phase high pressure and fast protein liquid), differential centrifugation, differential solubility, or by any other standard technique used for the purification of proteins. Functional properties can be evaluated using any suitable assay known in the art.

Alternatively, once an MTSP polypeptide or its domain or derivative is identified, the amino acid sequence of the protein can be deduced from the nucleotide sequence of the gene which encodes it. As a result, the protein or its domain or derivative can be synthesized by standard chemical methods known in the art (*e.g.* see Hunkapiller et al, *Nature* 310:105-111 (1984)).

Manipulations of MTSP polypeptide sequences can be made at the protein level. Also contemplated herein are MTSP polypeptide proteins, domains thereof, derivatives or analogs or fragments thereof, which are differentially modified during or after translation, *e.g.*, by glycosylation, acetylation, phosphorylation, amidation, derivatization by known protecting/blocking groups, proteolytic cleavage, linkage to an antibody molecule or other cellular ligand. Any of numerous chemical modifications can be carried out by known techniques, including but not limited to specific chemical cleavage by cyanogen bromide, trypsin, chymotrypsin, papain, V8 protease, NaBH₄, acetylation, formylation, oxidation, reduction, metabolic synthesis in the presence of tunicamycin and other such agents.

In addition, domains, analogs and derivatives of an MTSP polypeptide can be chemically synthesized. For example, a peptide corresponding to a portion of an MTSP polypeptide, which includes the desired domain or which mediates the desired activity *in vitro* can be synthesized by use of a peptide synthesizer. Furthermore, if desired, nonclassical amino acids or chemical amino acid analogs can be introduced as a substitution or addition into the MTSP polypeptide sequence. Non-classical amino acids include but are not limited to the D-isomers of the common amino acids, α-amino isobutyric acid, 4-aminobutyric acid, Abu,

-66-

2-aminobutyric acid, ϵ -Abu, ϵ -Ahx, 6-amino hexanoic acid, Aib, 2-amino isobutyric acid, 3-amino propionic acid, ornithine, norleucine, norvaline, hydroxyproline, sarcosine, citrulline, cysteic acid, t-butylglycine, t-butylalanine, phenylglycine, cyclohexylalanine, β -alanine, fluoro-amino acids, designer amino acids such as β -methyl amino acids, Ca-methyl amino acids, Na-methyl amino acids, and amino acid analogs in general. Furthermore, the amino acid can be D (dextrorotary) or L (levorotary).

In cases where natural products are suspected of being mutant or are isolated from new species, the amino acid sequence of the MTSP polypeptide isolated from the natural source, as well as those expressed *in vitro*, or from synthesized expression vectors *in vivo* or *in vitro*, can be determined from analysis of the DNA sequence, or alternatively, by direct sequencing of the isolated protein. Such analysis can be performed by manual sequencing or through use of an automated amino acid sequenator.

15 **Modifications**

A variety of modifications of the MTSP polypeptides and domains are contemplated herein. An MTSP-encoding nucleic acid molecule can be modified by any of numerous strategies known in the art (Sambrook *et al.* (1990), *Molecular Cloning, A Laboratory Manual*, 2d ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, New York). The sequences can be cleaved at appropriate sites with restriction endonuclease(s), followed by further enzymatic modification if desired, isolated, and ligated *in vitro*. In the production of the gene encoding a domain, derivative or analog of MTSP, care should be taken to ensure that the modified gene retains the original translational reading frame, uninterrupted by translational stop signals, in the gene region where the desired activity is encoded.

Additionally, the MTSP-encoding nucleic acid molecules can be mutated *in vitro* or *in vivo*, to create and/or destroy translation, initiation, and/or termination sequences, or to create variations in coding regions and/or form new restriction endonuclease sites or destroy pre-existing ones, to facilitate further *in vitro* modification. Also, as described herein muteins with primary sequence alterations, such as replacements of Cys residues and elimination or addition of

-67-

glycosylation sites are contemplated; the MTSP9 of SEQ ID No. 18 has two potential glycosylation sites. Such mutations can be effected by any technique for mutagenesis known in the art, including, but not limited to, chemical mutagenesis and *in vitro* site-directed mutagenesis (Hutchinson et al., *J. Biol.*

- 5 *Chem.* 253:6551-6558 (1978)), use of TAB® linkers (Pharmacia). In one embodiment, for example, an MTSP polypeptide or domain thereof is modified to include a fluorescent label. In other specific embodiments, the MTSP polypeptide is modified to have a heterobifunctional reagent, such
- 10 heterobifunctional reagents can be used to crosslink the members of the complex.

- In addition, domains, analogs and derivatives of an MTSP can be chemically synthesized. For example, a peptide corresponding to a portion of an MTSP, which includes the desired domain or which mediates the desired activity *in vitro* can be synthesized by use of a peptide synthesizer. Furthermore, if
- 15 desired, nonclassical amino acids or chemical amino acid analogs can be introduced as a substitution or addition into the MTSP sequence. Non-classical amino acids include but are not limited to the D-isomers of the common amino acids, α -amino isobutyric acid, 4-aminobutyric acid, Abu, 2-aminobutyric acid, ϵ -Abu, e-Ahx, 6-amino hexanoic acid, Aib, 2-amino isobutyric acid, 3-amino
- 20 propionic acid, ornithine, norleucine, norvaline, hydroxyproline, sarcosine, citrulline, cysteic acid, t-butylglycine, t-butylalanine, phenylglycine, cyclohexylalanine, β -alanine, fluoro-amino acids, designer amino acids such as β -methyl amino acids, Ca-methyl amino acids, Na-methyl amino acids, and amino acid analogs in general. Furthermore, the amino acid can be D (dextrorotary) or
- 25 L (levorotary).

F. SCREENING METHODS

- The single chain protease domains, as shown herein, can be used in a variety of methods to identify compounds that modulate the activity thereof. For
- 30 SPs that exhibit higher activity or expression in tumor cells, compounds that inhibit the proteolytic activity are of particular interest. For any SPs that are active at lower levels in tumor cells, compounds or agents that enhance the

-68-

activity are potentially of interest. In all instances the identified compounds include agents that are candidate cancer treatments.

Several types of assays are exemplified and described herein. It is understood that the protease domains can be used in other assays. It is shown here, however, that the single chain protease domains exhibit catalytic activity. As such they are ideal for *in vitro* screening assays. They can also be used in binding assays.

The MTSP9 full length zymogens, activated enzymes, single and two chain protease domains are contemplated for use in any screening assay known to those of skill in the art, including those provided herein. Hence the following description, if directed to proteolytic assays is intended to apply to use of a single chain protease domain or a catalytically active portion thereof of any SP, including an MTSP9. Other assays, such as binding assays are provided herein, particularly for use with an MTSP9, including any variants, such as splice variants thereof.

1. Catalytic Assays for identification of agents that modulate the protease activity of an SP protein

Methods for identifying a modulator of the catalytic activity of an SP, particularly a single chain protease domain or catalytically active portion thereof, are provided herein. The methods can be practiced by: contacting the MTSP9, a full-length zymogen or activated form, and particularly a single-chain domain thereof, with a substrate of the MTSP9 in the presence of a test substance, and detecting the proteolysis of the substrate, whereby the activity of the MTSP9 is assessed, and comparing the activity to a control. For example, a control can be the activity of the MTSP9 assessed by contacting an MTSP9, including a full-length zymogen or activated form, and particularly a single-chain domain thereof, particularly a single-chain domain thereof, with a substrate of the MTSP9, and detecting the proteolysis of the substrate, whereby the activity of the MTSP9 is assessed. The results in the presence and absence of the test compounds are compared. A difference in the activity indicates that the test substance modulates the activity of the MTSP9. Activators of MTSP9 activation cleavage are also contemplated; such assays are discussed below.

-69-

In one embodiment a plurality of the test substances are screened simultaneously in the above screening method. In another embodiment, the MTSP9 is isolated from a target cell as a means for then identifying agents that are potentially specific for the target cell.

- 5 In another embodiment, a test substance is a therapeutic compound, and whereby a difference of the MTSP9 activity measured in the presence and in the absence of the test substance indicates that the target cell responds to the therapeutic compound.

One method includes the steps of (a) contacting the MTSP9 polypeptide
10 or protease domain thereof with one or a plurality of test compounds under conditions conducive to interaction between the ligand and the compounds; and (b) identifying one or more compounds in the plurality that specifically binds to the ligand.

Another method provided herein includes the steps of a) contacting an
15 MTSP9 polypeptide or protease domain thereof with a substrate of the MTSP9 polypeptide, and detecting the proteolysis of the substrate, whereby the activity of the MTSP9 polypeptide is assessed; b) contacting the MTSP9 polypeptide with a substrate of the MTSP9 polypeptide in the presence of a test substance, and detecting the proteolysis of the substrate, whereby the activity of the
20 MTSP9 polypeptide is assessed; and c) comparing the activity of the MTSP9 polypeptide assessed in steps a) and b), whereby the activity measured in step a) differs from the activity measured in step b) indicates that the test substance modulates the activity of the MTSP9 polypeptide.

In another embodiment, a plurality of the test substances are screened
25 simultaneously. In comparing the activity of an MTSP9 polypeptide in the presence and absence of a test substance to assess whether the test substance is a modulator of the MTSP9 polypeptide, it is unnecessary to assay the activity in parallel, although such parallel measurement is typical. It is possible to measure the activity of the MTSP9 polypeptide at one time point and compare
30 the measured activity to a historical value of the activity of the MTSP9 polypeptide.

-70-

For instance, one can measure the activity of the MTSP9 polypeptide in the presence of a test substance and compare with historical value of the activity of the MTSP9 polypeptide measured previously in the absence of the test substance, and *vice versa*. This can be accomplished, for example, by
5 providing the activity of the MTSP9 polypeptide on an insert or pamphlet provided with a kit for conducting the assay.

Methods for selecting substrates for a particular SP are described in the EXAMPLES, and particular proteolytic assays are exemplified.

Combinations and kits containing the combinations optionally including
10 instructions for performing the assays are provided. The combinations include an MTSP9 polypeptide and a substrate of the MTSP9 polypeptide to be assayed; and, optionally reagents for detecting proteolysis of the substrate. The substrates, which can be chromogenic or fluorogenic molecules, including proteins, subject to proteolysis by a particular MTSP9 polypeptide, can be
15 identified empirically by testing the ability of the MTSP9 polypeptide to cleave the test substrate. Substrates that are cleaved most effectively (*i.e.*, at the lowest concentrations and/or fastest rate or under desirable conditions), are identified.

Additionally provided herein is a kit containing the above-described
20 combination. The kit optionally includes instructions for identifying a modulator of the activity of an MTSP9 polypeptide. Any MTSP9 polypeptide is contemplated as target for identifying modulators of the activity thereof.

2. Binding assays

Also provided herein are methods for identification and isolation of
25 agents, particularly compounds that bind to MTSP9s. The assays are designed to identify agents that bind to the zymogen form, the single chain isolated protease domain (or a protein, other than an MTSP9 polypeptide, that contains the protease domain of an MTSP9 polypeptide), and to the activated form, including the activated form derived from the full length zymogen or from an
30 extended protease domain. The identified compounds are candidates or leads for identification of compounds for treatments of tumors and other disorders and diseases involving aberrant angiogenesis. The MTSP9 polypeptides used in the

-71-

methods include any MTSP9 polypeptide as defined herein, including the MTSP9 single chain protease domain or proteolytically active portion thereof.

A variety of methods are provided herein. These methods can be performed in solution or in solid phase reactions in which the MTSP9 polypeptide(s) or protease domain(s) thereof are linked, either directly or indirectly via a linker, to a solid support. Screening assays are described in the Examples, and these assays have been used to identify candidate compounds. For purposes herein, all binding assays described above are provided for MTSP9.

Methods for identifying an agent, such as a compound, that specifically binds to an MTSP9 single chain protease domain, a zymogen or full-length activated MTSP9 or two chain protease domain thereof are provided herein. The method can be practiced by (a) contacting the MTSP9 with one or a plurality of test agents under conditions conducive to binding between the MTSP9 and an agent; and (b) identifying one or more agents within the plurality that specifically binds to the MTSP9.

For example, in practicing such methods the MTSP9 polypeptide is mixed with a potential binding partner or an extract or fraction of a cell under conditions that allow the association of potential binding partners with the polypeptide. After mixing, peptides, polypeptides, proteins or other molecules that have become associated with an MTSP9 are separated from the mixture. The binding partner that bound to the MTSP9 can then be removed and further analyzed. To identify and isolate a binding partner, the entire protein, for instance the entire disclosed protein of SEQ ID Nos. 6 can be used. Alternatively, a fragment of the protein can be used.

A variety of methods can be used to obtain cell extracts or body fluids, such as blood, serum, urine, sweat, synovial fluid, CSF and other such fluids. For example, cells can be disrupted using either physical or chemical disruption methods. Examples of physical disruption methods include, but are not limited to, sonication and mechanical shearing. Examples of chemical lysis methods include, but are not limited to, detergent lysis and enzyme lysis. A skilled artisan

-72-

can readily adapt methods for preparing cellular extracts in order to obtain extracts for use in the present methods.

Once an extract of a cell is prepared, the extract is mixed with the MTSP9 under conditions in which association of the protein with the binding partner can occur. A variety of conditions can be used, including conditions that resemble conditions found in the cytoplasm of a human cell or in a body fluid, such as blood. Features, such as osmolarity, pH, temperature, and the concentration of cellular extract used, can be varied to optimize the association of the protein with the binding partner. Similarly, methods for isolation of molecules of interest from body fluids are known.

After mixing under appropriate conditions, the bound complex is separated from the mixture. A variety of techniques can be used to separate the mixture. For example, antibodies specific to an MTSP9 can be used to immunoprecipitate the binding partner complex. Alternatively, standard chemical separation techniques such as chromatography and density/sediment centrifugation can be used.

After removing the non-associated cellular constituents in the extract, the binding partner can be dissociated from the complex using conventional methods. For example, dissociation can be accomplished by altering the salt concentration or pH of the mixture.

To aid in separating associated binding partner pairs from the mixed extract, the MTSP9 can be immobilized on a solid support. For example, the protein can be attached to a nitrocellulose matrix or acrylic beads. Attachment of the protein or a fragment thereof to a solid support aids in separating peptide/binding partner pairs from other constituents found in the extract. The identified binding partners can be either a single protein or a complex made up of two or more proteins.

Alternatively, the nucleic acid molecules encoding the single chain proteases can be used in a yeast two-hybrid system. The yeast two-hybrid system has been used to identify other protein partner pairs and can readily be adapted to employ the nucleic acid molecules herein described.

-73-

Another *in vitro* binding assay, particularly for an MTSP9, uses a mixture of a polypeptide that contains at least the catalytic domain of one of these proteins and one or more candidate binding targets or substrates. After incubating the mixture under appropriate conditions, the ability of the MTSP9 or a polypeptide fragment thereof containing the catalytic domain to bind to or interact with the candidate substrate is assessed. For cell-free binding assays, one of the components includes or is coupled to a detectable label. The label can provide for direct detection, such as radioactivity, luminescence, optical or electron density, *etc.*, or indirect detection such as an epitope tag, an enzyme, *etc.* A variety of methods can be employed to detect the label depending on the nature of the label and other assay components. For example, the label can be detected bound to the solid substrate or a portion of the bound complex containing the label can be separated from the solid substrate, and the label thereafter detected.

3. Detection of signal transduction

MTSP9, which is a transmembrane protein, can be involved directly or indirectly in signal transduction directly as a cell surface receptor or indirectly by activating proteins, such as pro-growth factors that can initiate signal transduction.

In addition, secreted of MTSP9, such as the extracellular domain of MTSP9, can be involved in signal transduction either directly by binding to or interacting with a cell surface receptor or indirectly by activating proteins, such as pro-growth factors that can initiate signal transduction. Assays for assessing signal transduction are well known to those of skill in the art, and can be adapted for use with the MTSP9 polypeptide.

Assays for identifying agents that affect or alter signal transduction mediated directly or indirectly, such as via activation of a pro-growth factor, by an MTSP9, particularly the full length or a sufficient portion to anchor the extracellular domain or a functional portion thereof of an MTSP9 on the surface of a cell are provided. Such assays, include, for example, transcription based assays in which modulation of a transduced signal is assessed by detecting an

-74-

effect on an expression from a reporter gene (see, *e.g.*, U.S. Patent No. 5,436,128).

4. Methods for Identifying Agents that Modulate the Expression a Nucleic Acid Encoding an MTSP9

5 Another embodiment provides methods for identifying agents that modulate the expression of a nucleic acid encoding an MTSP9. Such assays use any available means of monitoring for changes in the expression level of the nucleic acids encoding an MTSP9.

10 In one assay format, cell lines that contain reporter gene fusions between the open reading frame of MTSP9 or a domain thereof, particularly the protease domain and any assayable fusion partner can be prepared. Numerous assayable fusion partners are known and readily available including the firefly luciferase gene and the gene encoding chloramphenicol acetyltransferase (Alam *et al.*, *Anal. Biochem.* 188: 245-54 (1990)). Cell lines containing the reporter gene
15 fusions are then exposed to the agent to be tested under appropriate conditions and time. Differential expression of the reporter gene between samples exposed to the agent and control samples identifies agents which modulate the expression of a nucleic acid encoding an MTSP9.

20 Additional assay formats can be used to monitor the ability of the agent to modulate the expression of a nucleic acid encoding an MTSP9. For instance, mRNA expression can be monitored directly by hybridization to the nucleic acids. Cell lines are exposed to the agent to be tested under appropriate conditions and time and total RNA or mRNA is isolated by standard procedures (see, *e.g.*, Sambrook *et al.* (1989) MOLECULAR CLONING: A LABORATORY MANUAL,
25 2nd Ed. Cold Spring Harbor Laboratory Press). Probes to detect differences in RNA expression levels between cells exposed to the agent and control cells can be prepared from the nucleic acids. It is typical, but not necessary, to design probes which hybridize only with target nucleic acids under conditions of high stringency. Only highly complementary nucleic acid hybrids form under
30 conditions of high stringency. Accordingly, the stringency of the assay conditions determines the amount of complementarity which should exist between two nucleic acid strands in order to form a hybrid. Stringency should

-75-

be chosen to maximize the difference in stability between the probe:target hybrid and potential probe:non-target hybrids.

Probes can be designed from the nucleic acids through methods known in the art. For instance, the G + C content of the probe and the probe length can

5 affect probe binding to its target sequence. Methods to optimize probe specificity are commonly available (see, *e.g.*, Sambrook *et al.* (1989)

MOLECULAR CLONING: A LABORATORY MANUAL, 2nd Ed. Cold Spring Harbor Laboratory Press); and Ausubel *et al.* (1995) CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, Greene Publishing Co., NY).

10 Hybridization conditions are modified using known methods (see, *e.g.*, Sambrook *et al.* (1989) MOLECULAR CLONING: A LABORATORY MANUAL, 2nd Ed. Cold Spring Harbor Laboratory Press); and Ausubel *et al.* (1995) CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, Greene Publishing Co., NY), as required for each probe. Hybridization of total cellular RNA or RNA enriched
15 for polyA RNA can be accomplished in any available format. For instance, total cellular RNA or RNA enriched for polyA RNA can be affixed to a solid support, and the solid support exposed to at least one probe comprising at least one, or part of one of the nucleic acid molecules under conditions in which the probe specifically hybridizes. Alternatively, nucleic acid fragments comprising at least
20 one, or part of one of the sequences can be affixed to a solid support, such as a porous glass wafer. The glass wafer can then be exposed to total cellular RNA or polyA RNA from a sample under conditions in which the affixed sequences specifically hybridize. Such glass wafers and hybridization methods are widely available, for example, those disclosed by Beattie (WO 95/11755). By
25 examining for the ability of a given probe to specifically hybridize to an RNA sample from an untreated cell population and from a cell population exposed to the agent, agents which up or down regulate the expression of a nucleic acid encoding the MTSP9 polypeptide, are identified.

-76-

In one format, the relative amounts of a protein between a cell population that has been exposed to the agent to be tested compared to an un-exposed control cell population can be assayed (*e.g.*, a prostate cancer cell line, a lung cancer cell line, a colon cancer cell line or a breast cancer cell line). In this
5 format, probes, such as specific antibodies, are used to monitor the differential expression or level of activity of the protein in the different cell populations or body fluids. Cell lines or populations or body fluids are exposed to the agent to be tested under appropriate conditions and time. Cellular lysates or body fluids can be prepared from the exposed cell line or population and a control,
10 unexposed cell line or population or unexposed body fluid. The cellular lysates or body fluids are then analyzed with the probe.

For example, N- and C- terminal fragments of the MTSP9 can be expressed in bacteria and used to search for proteins which bind to these fragments. Fusion proteins, such as His-tag or GST fusion to the N- or C-
15 terminal regions of the MTSP9 can be prepared for use as a substrate. These fusion proteins can be coupled to, for example, Glutathione-Sepharose beads and then probed with cell lysates or body fluids. Prior to lysis, the cells or body fluids can be treated with a candidate agent which can modulate an MTSP9 or proteins that interact with domains thereon. Lysate proteins binding to the
20 fusion proteins can be resolved by SDS-PAGE, isolated and identified by protein sequencing or mass spectroscopy, as is known in the art.

Antibody probes are prepared by immunizing suitable mammalian hosts in appropriate immunization protocols using the peptides, polypeptides or proteins if they are of sufficient length (*e.g.*, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 20,
25 25, 30, 35, 40 or more consecutive amino acids the MTSP9 polypeptide or if required to enhance immunogenicity, conjugated to suitable carriers. Methods for preparing immunogenic conjugates with carriers, such as bovine serum albumin (BSA), keyhole limpet hemocyanin (KLH), or other carrier proteins are well known in the art. In some circumstances, direct conjugation using, for
30 example, carbodiimide reagents can be effective; in other instances linking reagents such as those supplied by Pierce Chemical Co., Rockford, IL, can be desirable to provide accessibility to the hapten. Hapten peptides can be

-77-

extended at either the amino or carboxy terminus with a Cys residue or interspersed with cysteine residues, for example, to facilitate linking to a carrier. Administration of the immunogens is conducted generally by injection over a suitable time period and with use of suitable adjuvants, as is generally

5 understood in the art. During the immunization schedule, titers of antibodies are taken to determine adequacy of antibody formation.

Anti-peptide antibodies can be generated using synthetic peptides corresponding to, for example, the carboxy terminal amino acids of the MTSP9. Synthetic peptides can be as small as 1-3 amino acids in length, generally at
10 least 4 or more amino acid residues long. The peptides can be coupled to KLH using standard methods and can be immunized into animals, such as rabbits or ungulates. Polyclonal antibodies can then be purified, for example using Actigel beads containing the covalently bound peptide.

While the polyclonal antisera produced in this way can be satisfactory for
15 some applications, for pharmaceutical compositions, use of monoclonal preparations are generally used. Immortalized cell lines which secrete the desired monoclonal antibodies can be prepared using the standard method of Kohler *et al.*, (*Nature* 256: 495-7 (1975)) or modifications which effect immortalization of lymphocytes or spleen cells, as is generally known. The
20 immortalized cell lines secreting the desired antibodies are screened by immunoassay in which the antigen is the peptide hapten, polypeptide or protein. When the appropriate immortalized cell culture secreting the desired antibody is identified, the cells can be cultured either *in vitro* or by production *in vivo* via ascites fluid. Of particular interest, are monoclonal antibodies that recognize the
25 catalytic domain or activation cleavage site (region) of an MTSP9.

Additionally, the zymogen or two-chain form of the MTSP9 can be used to make monoclonal antibodies that recognize conformation epitopes. The desired monoclonal antibodies are then recovered from the culture supernatant or from the ascites supernatant. Fragments of the monoclonals or the polyclonal
30 antisera which contain the immunologically significant portion can be used as antagonists, as well as the intact antibodies. Use of immunologically reactive fragments, such as the Fab, Fab', of F(ab')₂ fragments are often used, especially

-78-

in a therapeutic context, as these fragments are generally less immunogenic than the whole immunoglobulin.

The antibodies or fragments can also be produced. Regions that bind specifically to the desired regions of receptor also can be produced in the
5 context of chimeras with multiple species origin.

Agents that are assayed in the above method can be randomly selected or rationally selected or designed.

The agents can be, as examples, peptides, small molecules, and carbohydrates. A skilled artisan can readily recognize that there is no limit as to
10 the structural nature of the agents.

The peptide agents can be prepared using standard solid phase (or solution phase) peptide synthesis methods, as is known in the art. In addition, the DNA encoding these peptides can be synthesized using commercially available oligonucleotide synthesis instrumentation and produced recombinantly
15 using standard recombinant production systems. The production using solid phase peptide synthesis is necessitated if non-gene-encoded amino acids are to be included.

G. Assay formats and selection of test substances that modulate at least one activity of an MTSP9 polypeptide

20 Methods for identifying agents that modulate at least one activity of an MTSP9 are provided. The methods include phage display and other methods for assessing alterations in the activity of an MTSP9. Such methods or assays can use any means of monitoring or detecting the desired activity. A variety of formats and detection protocols are known for performing screening assays.
25 Any such formats and protocols can be adapted for identifying modulators of MTSP9 polypeptide activities. The following includes a discussion of exemplary protocols.

1. High throughput screening assays

Although the above-described assay can be conducted where a single
30 MTSP9 polypeptide is screened, and/or a single test substance is screened in one assay, the assay typically is conducted in a high throughput screening mode, *i.e.*, a plurality of the SP proteins are screened against and/or a plurality

-79-

of the test substances are screened simultaneously (*See generally, High Throughput Screening: The Discovery of Bioactive Substances* (Devlin, Ed.) Marcel Dekker, 1997; Sittampalam et al., *Curr. Opin. Chem. Biol.*, 1:384-91 (1997); and Silverman et al., *Curr. Opin. Chem. Biol.*, 2:397-403 (1998)). For
5 example, the assay can be conducted in a multi-well (*e.g.*, 24-, 48-, 96-, 384-, 1536-well or higher density), chip or array format.

High-throughput screening (HTS) is the process of testing a large number of diverse chemical structures against disease targets to identify "hits" (Sittampalam et al., *Curr. Opin. Chem. Biol.*, 1:384-91 (1997)). Current state-of-
10 the-art HTS operations are highly automated and computerized to handle sample preparation, assay procedures and the subsequent processing of large volumes of data.

Detection technologies employed in high-throughput screens depend on the type of biochemical pathway being investigated (Sittampalam et al., *Curr.*
15 *Opin. Chem. Biol.*, 1:384-91 (1997)). These methods include, radiochemical methods, such as the scintillation proximity assays (SPA), which can be adapted to a variety of enzyme assays (Lerner et al., *J. Biomol. Screening*, 1:135-143 (1996); Baker et al., *Anal. Biochem.*, 239:20-24 (1996); Baum et al., *Anal. Biochem.*, 237:129-134 (1996); and Sullivan et al., *J. Biomol. Screening* 2:19-
20 23 (1997)) and protein-protein interaction assays (Braunwalder et al., *J. Biomol. Screening* 1:23-26 (1996); Sonatore et al., *Anal. Biochem.* 240:289-297 (1996); and Chen et al., *J. Biol. Chem.* 271:25308-25315 (1996)), and non-isotopic detection methods, including but are not limited to, colorimetric and
luminescence detection methods, resonance energy transfer (RET) methods,
25 time-resolved fluorescence (HTRF) methods, cell-based fluorescence assays, such as fluorescence resonance energy transfer (FRET) procedures (*see, e.g.*, Gonzalez et al., *Biophys. J.*, 69:1272-1280 (1995)), fluorescence polarization or anisotropy methods (*see, e.g.*, Jameson et al., *Methods Enzymol.* 246:283-300 (1995); Jolley, *J. Biomol. Screening* 1:33-38 (1996); Lynch et al.,
30 *Anal. Biochem.* 247:77-82 (1997)), fluorescence correlation spectroscopy (FCS) and other such methods.

-80-

2. Test Substances

Test compounds, including small molecules, antibodies, proteins, nucleic acids, peptides, and libraries and collections thereof, can be screened in the above-described assays and assays described below to identify compounds that modulate the activity of an MTSP9 polypeptide. Rational drug design methodologies that rely on computational chemistry can be used to screen and identify candidate compounds.

The compounds identified by the screening methods include inhibitors, including antagonists, and can be agonists. Compounds for screening include any compounds and collections of compounds available, known or that can be prepared.

a. Selection of Compounds

Compounds can be selected for their potency and selectivity of inhibition of serine proteases, especially an MTSP9 polypeptide. As described herein, and as generally known, a target serine protease and its substrate are combined under assay conditions permitting reaction of the protease with its substrate. The assay is performed in the absence of test compound, and in the presence of increasing concentrations of the test compound. The concentration of test compound at which 50% of the serine protease activity is inhibited by the test compound is the IC_{50} value (Inhibitory Concentration) or EC_{50} (Effective Concentration) value for that compound. Within a series or group of test compounds, those having lower IC_{50} or EC_{50} values are considered more potent inhibitors of the serine protease than those compounds having higher IC_{50} or EC_{50} values. The IC_{50} measurement is often used for more simplistic assays, whereas the EC_{50} is often used for more complicated assays, such as those employing cells.

Typically candidate compounds have an IC_{50} value of 100 nM or less as measured in an *in vitro* assay for inhibition of MTSP9 polypeptide activity. The test compounds also are evaluated for selectivity toward a serine protease. As described herein, and as generally known, a test compound is assayed for its potency toward a panel of serine proteases and other enzymes and an IC_{50} value or EC_{50} value is determined for each test compound in each assay system. A

-81-

compound that demonstrates a low IC_{50} value or EC_{50} value for the target enzyme, *e.g.*, MTSP9 polypeptide, and a higher IC_{50} value or EC_{50} value for other enzymes within the test panel (*e.g.*, urokinase tissue plasminogen activator, thrombin, Factor Xa), is considered to be selective toward the target enzyme.

- 5 Generally, a compound is deemed selective if its IC_{50} value or EC_{50} value in the target enzyme assay is at least one order of magnitude less than the next smallest IC_{50} value or EC_{50} value measured in the selectivity panel of enzymes.

Compounds are also evaluated for their activity *in vivo*. The type of assay chosen for evaluation of test compounds depends on the pathological
10 condition to be treated or prevented by use of the compound, as well as the route of administration to be evaluated for the test compound.

For instance, to evaluate the activity of a compound to reduce tumor growth through inhibition of MTSP9 polypeptide, the procedures described by Jankun et al., *Canc. Res.* 57:559-563 (1997) to evaluate PAI-1 can be
15 employed. Briefly, the ATCC cell lines DU145 and LnCaP are injected into SCID mice. After tumors are established, the mice are given test compound according to a dosing regime determined from the compound's *in vitro* characteristics. The Jankun *et al.* compound was administered in water. Tumor volume measurements are taken twice a week for about five weeks. A compound is
20 deemed active if an animal to which the compound was administered exhibited decreased tumor volume, as compared to animals receiving appropriate control compounds.

Another *in vivo* experimental model designed to evaluate the effect of p-aminobenzamidine, a swine protease inhibitor, on reducing tumor volume is
25 described by Billström et al., *Int. J. Cancer* 61:542-547 (1995).

To evaluate the ability of a compound to reduce the occurrence of, or inhibit, metastasis, the procedures described by Kobayashi *et al.* *Int. J. Canc.* 57:727-733d (1994) can be employed. Briefly, a murine xenograft selected for high lung colonization potential is injected into C57B1/6 mice i.v. (experimental
30 metastasis) or s.c. into the abdominal wall (spontaneous metastasis). Various concentrations of the compound to be tested can be admixed with the tumor cells in Matrigel prior to injection. Daily i.p. injections of the test compound are

-82-

made either on days 1-6 or days 7-13 after tumor inoculation. The animals are sacrificed about three or four weeks after tumor inoculation, and the lung tumor colonies are counted. Evaluation of the resulting data permits a determination as to efficacy of the test compound, optimal dosing and route of administration.

5 The activity of the tested compounds toward decreasing tumor volume and metastasis can be evaluated in model described in Rabbani et al., *Int. J. Cancer* 63:840-845 (1995) to evaluate their inhibitor. There, Mat LyLu tumor cells were injected into the flank of Copenhagen rats. The animals were implanted with osmotic minipumps to continuously administer various doses of
10 test compound for up to three weeks. The tumor mass and volume of experimental and control animals were evaluated during the experiment, as were metastatic growths. Evaluation of the resulting data permits a determination as to efficacy of the test compound, optimal dosing, and route of administration. Some of these authors described a related protocol in Xing et al., *Canc. Res.*
15 57:3585-3593 (1997).

To evaluate the anti-angiogenesis activity of a compound, a rabbit cornea neovascularization model can be employed (see, e.g., Avery et al. (1990) *Arch. Ophthalmol.*, 108:1474-147). Avery et al. describes anesthetizing New Zealand albino rabbits and then making a central corneal incision and forming a radial
20 corneal pocket. A slow release prostaglandin pellet was placed in the pocket to induce neovascularization. Test compound was administered i.p. for five days, at which time the animals were sacrificed. The effect of the test compound is evaluated by review of periodic photographs taken of the limbus, which can be used to calculate the area of neovascular response and, therefore, limbal
25 neovascularization. A decreased area of neovascularization as compared with appropriate controls indicates the test compound was effective at decreasing or inhibiting neovascularization.

An angiogenesis model used to evaluate the effect of a test compound in preventing angiogenesis is described by Min et al. *Canc. Res.* 56:2428-2433
30 (1996). C57BL6 mice receive subcutaneous injections of a Matrigel mixture containing bFGF, as the angiogenesis-inducing agent, with and without the test compound. After five days, the animals are sacrificed and the Matrigel plugs, in

-83-

which neovascularization can be visualized, are photographed. An experimental animal receiving Matrigel and an effective dose of test compound exhibits less vascularization than a control animal or an experimental animal receiving a less- or non-effective dose of compound.

5 An *in vivo* system designed to test compounds for their ability to limit the spread of primary tumors is described by Crowley et al., *Proc. Natl. Acad. Sci.* 90:5021-5025 (1993). Nude mice are injected with tumor cells (PC3) engineered to express CAT (chloramphenicol acetyltransferase). Compounds to be tested for their ability to decrease tumor size and/or metastases are
10 administered to the animals, and subsequent measurements of tumor size and/or metastatic growths are made. In addition, the level of CAT detected in various organs provides an indication of the ability of the test compound to inhibit metastasis; detection of less CAT in tissues of a treated animal versus a control animal indicates less CAT-expressing cells migrated to that tissue.

15 *In vivo* experimental modes designed to evaluate the inhibitory potential of a test serine protease inhibitors, using a tumor cell line F3II known to be highly invasive (see, e.g., Alonso et al., *Breast Canc. Res. Treat.* 40:209-223 (1996)). Alonso describes *in vivo* studies for toxicity determination, tumor growth, invasiveness, spontaneous metastasis, experimental lung metastasis,
20 and an angiogenesis assay.

 The CAM model (chick embryo chorioallantoic membrane model), first described by L. Ossowski in 1998 (*J. Cell Biol.* 107:2437-2445 (1988)), provides another method for evaluating the inhibitory activity of a test compound. In the CAM model, tumor cells invade through the chorioallantoic
25 membrane containing CAM (with tumor cells in the presence of several serine protease inhibitors results in less or no invasion of the tumor cells through the membrane). Thus, the CAM assay is performed with CAM and tumor cells in the presence and absence of various concentrations of test compound. The invasiveness of tumor cells is measured under such conditions to provide an
30 indication of the compound's inhibitory activity. A compound having inhibitory activity correlates with less tumor invasion.

-84-

The CAM model is also used in a standard assay of angiogenesis (*i.e.*, effect on formation of new blood vessels (Brooks *et al. Methods in Molecular Biology* 129:257-269 (1999)). According to this model, a filter disc containing an angiogenesis inducer, such as basic fibroblast growth factor (bFGF) is placed
5 onto the CAM. Diffusion of the cytokine into the CAM induces local angiogenesis, which can be measured in several ways such as by counting the number of blood vessel branch points within the CAM directly below the filter disc. The ability of identified compounds to inhibit cytokine-induced angiogenesis can be tested using this model. A test compound can either be
10 added to the filter disc that contains the angiogenesis inducer, be placed directly on the membrane or be administered systemically. The extent of new blood vessel formation in the presence and/or absence of test compound can be compared using this model. The formation of fewer new blood vessels in the presence of a test compound would be indicative of anti-angiogenesis activity.
15 Demonstration of anti-angiogenesis activity for inhibitors of an MTSP9 polypeptide indicates a role in angiogenesis for that SP protein.

b. Known serine protease inhibitors

Compounds for screening can be serine protease inhibitors, which can be tested for their ability to inhibit the activity of an MTSP9.
20 Exemplary, serine protease inhibitors for use in the screening assays, include, but are not limited to: Serine Protease Inhibitor 3 (SPI-3) (Chen, *et al. Citokine*, 11:856-862 (1999)); Aprotinin (Iijima, R., *et al., J. Biochem. (Tokyo)* 126:912-916 (1999)); Kazal-type serine protease inhibitor-like proteins (Niimi, *et al. Eur. J. Biochem.*, 266:282-292 (1999)); Kunitz-type serine protease inhibitor
25 (Ravichandran, S., *et al., Acta Crystallogr. D. Biol. Crystallogr.*, 55:1814-1821 (1999)); Tissue factor pathway inhibitor-2/Matrix-associated serine protease inhibitor (TFPI-2/MSPI), (Liu, Y. *et al. Arch. Biochem. Biophys.* 370:112-8 (1999)); Bukunin (Cui, C.Y. *et al. J. Invest. Dermatol.* 113:182-8 (1999)); Nafmostat mesilate (Ryo, R. *et al. Vox Sang.* 76:241-6 (1999)); TPCK (Huang
30 *et al. Oncogene* 18:3431-3439 (1999)); A synthetic cotton-bound serine protease inhibitor (Edwards *et al. Wound Repair Regen.* 7:106-18 (1999)); FUT-175 (Sawada, M. *et al. Stroke* 30:644-50 (1999)); Combination of serine

protease inhibitor FUT-0175 and thromboxane synthetase inhibitor OKY-046 (Kaminogo *et al. Neurol. Med. Chir. (Tokyo)* 38:704-8; discussion 708-9 (1998)); The rat serine protease inhibitor 2.1 gene (LeCam, A., et al., *Biochem. Biophys. Res. Commun.*, 253:311-4 (1998)); A new intracellular serine protease inhibitor expressed in the rat pituitary gland complexes with granzyme B (Hill *et al. FEBS Lett.* 440:361-4 (1998)); 3,4-Dichloroisocoumarin (Hammed *et al. Proc. Soc. Exp. Biol. Med.*, 219:132-7 (1998)); LEX032 (Bains *et al. Eur. J. Pharmacol.* 356:67-72 (1998)); N-tosyl-L-phenylalanine chloromethyl ketone (Dryjanski *et al. Biochemistry* 37:14151-6 (1998)); Mouse gene for the serine protease inhibitor neuroserpin (P112) (Berger *et al. Gene*, 214:25-33 (1998)); Rat serine protease inhibitor 2.3 gene (Paul *et al. Eur. J. Biochem.* 254:538-46 (1998)); Ecotin (Yang *et al. J. Mol. Biol.* 279:945-57 (1998)); A 14 kDa plant-related serine protease inhibitor (Roch *et al. Dev. Comp. Immunol.* 22(1):1-12 (1998)); Matrix-associated serine protease inhibitor TFPI-2/33 kDa MSPI (Rao *et al. Int. J. Cancer* 76:749-56 (1998)); ONO-3403 (Hiwasa *et al. Cancer Lett.* 126:221-5 (1998)); Bdellastasin (Moser *et al. Eur. J. Biochem.* 253:212-20 (1998)); Bikunin (Xu *et al. J. Mol. Biol.* 276:955-66 (1998)); Nafamostat mesilate (Mellgren *et al. Thromb. Haemost.* 79:342-7 (1998)); The growth hormone dependent serine protease inhibitor, Spi 2.1 (Maake *et al. Endocrinology* 138:5630-6 (1997)); Growth factor activator inhibitor type 2, a Kunitz-type serine protease inhibitor (Kawaguchi *et al. J. Biol. Chem.*, 272:27558-64 (1997)); Heat-stable serine protease inhibitor protein from ovaries of the desert locust, *Schistocerca gregaria* (Hamdaoui *et al. Biochem. Biophys. Res. Commun.* 238:357-60 (1997)); Human placental Hepatocyte growth factor activator inhibitor, a Kunitz-type serine protease inhibitor (Shimomura *et al. J. Biol. Chem.* 272:6370-6 (1997)); FUT-187, oral serine protease inhibitor (Shiozaki *et al. Gan To Kagaku Ryoho*, 23(14): 1971-9 (1996)); Extracellular matrix-associated serine protease inhibitors (Mr 33,000, 31,000, and 27,000 (Rao, C.N., et al., *Arch. Biochem. Biophys.*, 335:82-92 (1996)); An irreversible isocoumarin serine protease inhibitor (Palencia, D.D., et al., *Biol. Reprod.*, 55:536-42 (1996)); 4-(2-aminoethyl)-benzenesulfonyl fluoride (AEBSF) (Nakabo *et al. J. Leukoc. Biol.* 60:328-36 (1996)); Neuroserpin (Osterwalder, T., et al.,

- EMBO J.* 15:2944-53 (1996)); Human serine protease inhibitor alpha-1-antitrypsin (Forney *et al. J. Parasitol.* 82:496-502 (1996)); Rat serine protease inhibitor 2.3 (Simar-Blanchet, A.E., *et al., Eur. J. Biochem.*, 236:638-48 (1996)); Gebaxate mesilate (parodi, F., *et al., J. Cardiothorac. Vasc. Anesth.* 10:235-7 (1996));
- 5 Recombinant serine protease inhibitor, CPTI II (Stankiewicz, M., *et al., (Acta Biochim. Pol.*, 43(3):525-9 (1996)); A cysteine-rich serine protease inhibitor (Guamerin II) (Kim, D.R., *et al., J. Enzym. Inhib.*, 10:81-91 (1996)); Diisopropylfluorophosphate (Lundqvist, H., *et al., Inflamm. Res.*, 44(12):510-7 (1995)); Nexin 1 (Yu, D.W., *et al., J. Cell Sci.*, 108(Pt 12):3867-74 (1995));
- 10 LEX032 (Scalia, R., *et al., Shock*, 4(4):251-6 (1995)); Protease nexin I (Houenou, L.J., *et al., Proc. Natl. Acad. Sci. U.S.A.*, 92(3):895-9 (1995)); Chymase-directed serine protease inhibitor (Woodard S.L., *et al., J. Immunol.*, 153(11):5016-25 (1994)); N-alpha-tosyl-L-lysyl-chloromethyl ketone (TLCK) (Bourinbaier, A.S., *et al., Cell Immunol.*, 155(1):230-6 (1994)); Smpi56
- 15 (Ghendler, Y., *et al., Exp. Parasitol.*, 78(2):121-31 (1994)); *Schistosoma haematobium* serine protease (Blanton, R.E., *et al., Mol. Biochem. Parasitol.*, 63(1):1-11 (1994)); Spi-1 (Warren, W.C., *et al., Mol. Cell Endocrinol.*, 98(1):27-32 (1993)); TAME (Jessop, J.J., *et al., Inflammation*, 17(5):613-31 (1993)); Antithrombin III (Kalaria, R.N., *et al., Am. J. Pathol.*, 143(3):886-93 (1993));
- 20 FOY-305 (Ohkoshi, M., *et al., Anticancer Res.*, 13(4):963-6 (1993)); Camostat mesilate (Senda, S., *et al., Intern. Med.*, 32(4):350-4 (1993)); Pigment epithelium-derived factor (Steele, F.R., *et al., Proc. Natl. Acad. Sci. U.S.A.*, 90(4):1526-30 (1993)); Antistasin (Holstein, T.W., *et al., FEBS Lett.*, 309(3):288-92 (1992)); The vaccinia virus K2L gene encodes a serine protease
- 25 inhibitor (Zhou, J., *et al., Virology*, 189(2):678-86 (1992)); Bowman-Birk serine-protease inhibitor (Werner, M.H., *et al., J. Mol. Biol.*, 225(3):873-89 (1992); FUT-175 (Yanamoto, H., *et al., Neurosurgery*, 30(3):358-63 (1992)); FUT-175; (Yanamoto, H., *et al., Neurosurgery*, 30(3):351-6, discussion 356-7 (1992)); PAI-I (Yreadwell, B.V., *et al., J. Orthop. Res.*, 9(3):309-16 (1991)); 3,4-
- 30 Dichloroisocoumarin (Rusbridge, N.M., *et al., FEBS Lett.*, 268(1):133-6 (1990)); Alpha 1-antichymotrypsin (Lindmark, B.E., *et al., Am. Rev. Respir. Des.*, 141(4 Pt 1):884-8 (1990)); P-toluenesulfonyl-L-arginine methyl ester (TAME) (Scuderi,

-87-

- P., *J. Immunol.*, 143(1):168-73 (1989)); Alpha 1-antichymotrypsin (Abraham, C.R., et al., *Cell*, 52(4):487-501 (1988)); Contrapsin (Modha, J., et al., *Parasitology*, 96 (Pt 1):99-109 (1988)); Alpha 2-antiplasmin (Holmes, W.E., et al., *J. Biol. Chem.*, 262(4):1659-64 (1987)); 3,4-dichloroisocoumarin (Harper, J.W., et al., *Biochemistry*, 24(8):1831-41 (1985)); Diisopropylfluorophosphate (Tsutsui, K., et al., *Biochem. Biophys. Res. Commun.*, 123(1):271-7 (1984)); Gabexate mesilate (Hesse, B., et al., *Pharmacol. Res. Commun.*, 16(7):637-45 (1984)); Phenyl methyl sulfonyl fluoride (Dufer, J., et al., *Scand. J. Haematol.*, 32(1):25-32 (1984)); Protease inhibitor CI-2 (McPhalen, C.A., et al., *J. Mol. Biol.*, 168(2):445-7 (1983)); Phenylmethylsulfonyl fluoride (Sekar V., et al., *Biochem. Biophys. Res. Commun.*, 89(2):474-8 (1979)); PGE1 (Feinstein, M.D., et al., *Prostaglandine*, 14(6):1075-93 (1977)).

c. Combinatorial libraries and other libraries

- The source of compounds for the screening assays, can be libraries, including, but are not limited to, combinatorial libraries. Methods for synthesizing combinatorial libraries and characteristics of such combinatorial libraries are known in the art (*See generally, Combinatorial Libraries: Synthesis, Screening and Application Potential* (Cortese Ed.) Walter de Gruyter, Inc., 1995; Tietze and Lieb, *Curr. Opin. Chem. Biol.*, 2(3):363-71 (1998); Lam, *Anticancer Drug Des.*, 12(3):145-67 (1997); Blaney and Martin, *Curr. Opin. Chem. Biol.*, 1(1):54-9 (1997); and Schultz and Schultz, *Biotechnol. Prog.*, 12(6):729-43 (1996)).

- Methods and strategies for generating diverse libraries, primarily peptide- and nucleotide-based oligomer libraries, have been developed using molecular biology methods and/or simultaneous chemical synthesis methodologies (*see, e.g.,* Dower et al., *Annu. Rep. Med. Chem.*, 26:271-280 (1991); Fodor et al., *Science*, 251:767-773 (1991); Jung et al., *Angew. Chem. Ind. Ed. Engl.*, 31:367-383 (1992); Zuckerman et al., *Proc. Natl. Acad. Sci. USA*, 89:4505-4509 (1992); Scott et al., *Science*, 249:386-390 (1990); Devlin et al., *Science*, 249:404-406 (1990); Cwirla et al., *Proc. Natl. Acad. Sci. USA*, 87:6378-6382 (1990); and Gallop et al., *J. Medicinal Chemistry*, 37:1233-1251 (1994)). The

-88-

resulting combinatorial libraries potentially contain millions of compounds and that can be screened to identify compounds that exhibit a selected activity.

The libraries fall into roughly three categories: fusion-protein-displayed peptide libraries in which random peptides or proteins are presented on the surface of phage particles or proteins expressed from plasmids; support-bound synthetic chemical libraries in which individual compounds or mixtures of compounds are presented on insoluble matrices, such as resin beads (*see, e.g.,* Lam et al., *Nature*, 354:82-84 (1991)) and cotton supports (*see, e.g.,* Eichler et al., *Biochemistry* 32:11035-11041 (1993)); and methods in which the compounds are used in solution (*see, e.g.,* Houghten et al., *Nature*, 354:84-86 (1991); Houghten et al., *BioTechniques*, 313:412-421 (1992); and Scott et al., *Curr. Opin. Biotechnol.*, 5:40-48 (1994)). There are numerous examples of synthetic peptide and oligonucleotide combinatorial libraries and there are many methods for producing libraries that contain non-peptidic small organic molecules. Such libraries can be based on a basis set of monomers that are combined to form mixtures of diverse organic molecules or that can be combined to form a library based upon a selected pharmacophore monomer.

Either a random or a deterministic combinatorial library can be screened by the presently disclosed and/or claimed screening methods. In either of these two libraries, each unit of the library is isolated and/or immobilized on a solid support. In the deterministic library, one knows *a priori* a particular unit's location on each solid support. In a random library, the location of a particular unit is not known *a priori* although each site still contains a single unique unit. Many methods for preparing libraries are known to those of skill in this art (*see, e.g.,* Geysen et al., *Proc. Natl. Acad. Sci. USA*, 81:3998-4002 (1984), Houghten et al., *Proc. Natl. Acad. Sci. USA*, 81:5131-5135 (1985)). Combinatorial library generated by the any techniques known to those of skill in the art are contemplated (*see, e.g.,* Table 1 of Schultz and Schultz, *Biotechnol. Prog.*, 12(6):729-43 (1996)) for screening; Bartel et al., *Science*, 261:1411-1418 (1993); Baumbach et al. *BioPharm*, (Can):24-35 (1992); Bock et al. *Nature*, 355:564-566 (1992); Borman, S., Combinatorial chemists focus on small molecules molecular recognition, and automation, *Chem. Eng. News*,

- 2(12):29 (1996); Boublik, et al., Eukaryotic Virus Display: Engineering the Major Surface Glycoproteins of the Autographa California Nuclear Polyhedrosis Virus (ACNPV) for the Presentation of Foreign Proteins on the Virus Surface, *Bio/Technology*, 13:1079-1084 (1995); Brenner, et al., Encoded Combinatorial
- 5 Chemistry, *Proc. Natl. Acad. Sci. U.S.A.*, 89:5381-5383 (1992); Caflisch, et al., Computational Combinatorial Chemistry for *De Novo* Ligand Design: Review and Assessment, *Perspect. Drug Discovery Des.*, 3:51-84 (1995); Cheng, et al., Sequence-Selective Peptide Binding with a Peptido-A,B-*trans*-steroidal Receptor Selected from an Encoded Combinatorial Library, *J. Am. Chem. Soc.*, 118:1813-
- 10 1814 (1996); Chu, et al., Affinity Capillary Electrophoresis to Identify the Peptide in A Peptide Library that Binds Most Tightly to Vancomycin, *J. Org. Chem.*, 58:648-652 (1993); Clackson, et al., Making Antibody Fragments Using Phage Display Libraries, *Nature*, 352:624-628 (1991); Combs, et al., Protein Structure-Based Combinatorial Chemistry: Discovery of Non-Peptide Binding
- 15 Elements to Src SH3 Domain, *J. Am. Chem. Soc.*, 118:287-288 (1996); Cwirla, et al., Peptides On Phage: A Vast Library of Peptides for Identifying Ligands, *Proc. Natl. Acad. Sci. U.S.A.*, 87:6378-6382 (1990); Ecker, et al., Combinatorial Drug Discovery: Which Method will Produce the Greatest Value, *Bio/Technology*, 13:351-360 (1995); Ellington, et al., *In Vitro* Selection of RNA
- 20 Molecules That Bind Specific Ligands, *Nature*, 346:818-822 (1990); Ellman, J.A., Variants of Benzodiazepines, *J. Am. Chem. Soc.*, 114:10997 (1992); Erickson, et al., *The Proteins*; Neurath, H., Hill, R.L., Eds.: Academic: New York, 1976; pp. 255-257; Felici, et al., *J. Mol. Biol.*, 222:301-310 (1991); Fodor, et al., Light-Directed, Spatially Addressable Parallel Chemical Synthesis, *Science*,
- 25 251:767-773 (1991); Francisco, et al., Transport and Anchoring of Beta-Lactamase to the External Surface of *E. Coli.*, *Proc. Natl. Acad. Sci. U.S.A.*, 89:2713-2717 (1992); Georgiou, et al., Practical Applications of Engineering Gram-Negative Bacterial Cell Surfaces, *TIBTECH*, 11:6-10 (1993); Geysen, et al., Use of peptide synthesis to probe viral antigens for epitopes to a resolution of a
- 30 single amino acid, *Proc. Natl. Acad. Sci. U.S.A.*, 81:3998-4002 (1984); Glaser, et al., Antibody Engineering by Condon-Based Mutagenesis in a Filamentous Phage Vector System, *J. Immunol.*, 149:3903-3913 (1992); Gram, et al., In

- vitro selection and affinity maturation of antibodies from a naive combinatorial immunoglobulin library, *Proc. Natl. Acad. Sci.*, 89:3576-3580 (1992); Han, et al., Liquid-Phase Combinatorial Synthesis, *Proc. Natl. Acad. Sci. U.S.A.*, 92:6419-6423 (1995); Hoogenboom, et al., Multi-Subunit Proteins on the
- 5 Surface of Filamentous Phage: Methodologies for Displaying Antibody (Fab) Heavy and Light Chains, *Nucleic Acids Res.*, 19:4133-4137 (1991); Houghten, et al., General Method for the Rapid Solid-Phase Synthesis of Large Numbers of Peptides: Specificity of Antigen-Antibody Interaction at the Level of Individual Amino Acids, *Proc. Natl. Acad. Sci. U.S.A.*, 82:5131-5135 (1985); Houghten,
- 10 et al., The Use of Synthetic Peptide Combinatorial Libraries for the Determination of Peptide Ligands in Radio-Receptor Assays-Opioid-Peptides, *Bioorg. Med. Chem. Lett.*, 3:405-412 (1993); Houghten, et al., Generation and Use of Synthetic Peptide Combinatorial Libraries for Basic Research and Drug Discovery, *Nature*, 354:84-86 (1991); Huang, et al., Discovery of New Ligand Binding
- 15 Pathways in Myoglobin by Random Mutagenesis, *Nature Struct. Biol.*, 1:226-229 (1994); Huse, et al., Generation of a Large Combinatorial Library of the Immunoglobulin Repertoire In Phage Lambda, *Science*, 246:1275-1281 (1989); Janda, K.D., New Strategies for the Design of Catalytic Antibodies, *Biotechnol. Prog.*, 6:178-181 (1990); Jung, et al., Multiple Peptide Synthesis Methods and
- 20 Their Applications, *Angew. Chem. Int. Ed. Engl.*, 31:367-486 (1992); Kang, et al., Linkage of Recognition and Replication Functions By Assembling Combinatorial Antibody Fab Libraries Along Phage Surfaces, *Proc. Natl. Acad. Sci. U.S.A.*, 88:4363-4366 (1991a); Kang, et al., Antibody Redesign by Chain Shuffling from Random Combinatorial Immunoglobulin Libraries, *Proc. Natl.*
- 25 *Acad. Sci. U.S.A.*, 88:11120-11123 (1991b); Kay, et al., An M13 Phage Library Displaying Random 38-Amino-Acid-Peptides as a Source of Novel Sequences with Affinity to Selected Targets Genes, *Gene*, 128:59-65 (1993); Lam, et al., A new type of synthetic peptide library for identifying ligand-binding activity, *Nature*, 354:82-84 (1991) (published errata appear in *Nature*, 358:434 (1992)
- 30 and *Nature*, 360:768 (1992); Lebl, et al., One Bead One Structure Combinatorial Libraries, *Biopolymers (Pept. Sci.)*, 37:177-198 (1995); Lerner, et al., Antibodies without Immunization, *Science*, 258:1313-1314 (1992); Li, et al., Minimization

- of a Polypeptide Hormone, *Science*, 270:1657-1660 (1995); Light, et al., Display of Dimeric Bacterial Alkaline Phosphatase on the Major Coat Protein of Filamentous Bacteriophage, *Bioorg. Med. Chem. Lett.*, 3:1073-1079 (1992); Little, et al., Bacterial Surface Presentation of Proteins and Peptides: An
- 5 Alternative to Phage Technology, *Trends Biotechnol.*, 11:3-5 (1993); Marks, et al., By-Passing Immunization. Human Antibodies from V-Gene Libraries Displayed on Phage, *J. Mol. Biol.*, 222:581-597 (1991); Matthews, et al., Substrate Phage: Selection of Protease Substrates by Monovalent Phage Display, *Science*, 260:1113-1117 (1993); McCafferty, et al., Phage Enzymes: Expression
- 10 and Affinity Chromatography of Functional Alkaline Phosphatase on the Surface of Bacteriophage, *Protein Eng.*, 4:955-961 (1991); Menger, et al., Phosphatase Catalysis Developed Via Combinatorial Organic Chemistry, *J. Org. Chem.*, 60:6666-6667 (1995); Nicolaou, et al., *Angew. Chem. Int. Ed. Engl.*, 34:2289-2291 (1995); Oldenburg, et al., Peptide Ligands for A Sugar-Binding Protein
- 15 Isolated from a Random Peptide Library, *Proc. Natl. Acad. Sci. U.S.A.*, 89:5393-5397 (1992); Parmley, et al., Antibody-Selectable Filamentous fd Phage Vectors: Affinity Purification of Target Genes, *Genes*, 73:305-318 (1988); Pinilla, et al., Synthetic Peptide Combinatorial Libraries (SPCLS)--Identification of the Antigenic Determinant of Beta-Endorphin Recognized by Monoclonal Antibody-3E7, *Gene*,
- 20 128:71-76 (1993); Pinilla, et al., Review of the Utility of Soluble Combinatorial Libraries, *Biopolymers*, 37:221-240 (1995); Pistor, et al., Expression of Viral Hemagglutinin On the Surface of *E. Coli*, *Klin. Wochenschr.*, 66:110-116 (1989); Pollack, et al., Selective Chemical Catalysis by an Antibody, *Science*, 234:1570-1572 (1986); Rigler, et al., Fluorescence Correlations, Single Molecule
- 25 Detection and Large Number Screening: Applications in Biotechnology, *J. Biotechnol.*, 41:177-186 (1995); Sarvetnick, et al., Increasing the Chemical Potential of the Germ-Line Antibody Repertoire, *Proc. Natl. Acad. Sci. U.S.A.*, 90:4008-4011 (1993); Sastry, et al., Cloning of the Immunological Repertoire in *Escherichia Coli* for Generation of Monoclonal Catalytic Antibodies: Construction
- 30 of a Heavy Chain Variable Region-Specific cDNA Library, *Proc. Natl. Acad. Sci. U.S.A.*, 86:5728-5732 (1989); Scott, et al., Searching for Peptide Ligands with an Epitope Library, *Science*, 249:386-390 (1990); Sears, et al., Engineering

- Enzymes for Bioorganic Synthesis: Peptide Bond Formation, *Biotechnol. Prog.*, 12:423-433 (1996); Simon, et. al., Peptides: A Modular Approach to Drug Discovery, *Proc. Natl. Acad. Sci. U.S.A.*, 89:9367-9371 (1992); Still, et al., Discovery of Sequence-Selective Peptide Binding by Synthetic Receptors Using
- 5 Encoded Combinatorial Libraries, *Acc. Chem. Res.*, 29:155-163 (1996); Thompson, et al., Synthesis and Applications of Small Molecule Libraries, *Chem. Rev.*, 96:555-600 (1996); Tramontano, et al., Catalytic Antibodies, *Science*, 234:1566-1570 (1986); Wrighton, et al., Small Peptides as Potent Mimetics of the Protein Hormone Erythropoietin, *Science*, 273:458-464 (1996); York, et al.,
- 10 Combinatorial mutagenesis of the reactive site region in plasminogen activator inhibitor I, *J. Biol. Chem.*, 266:8595-8600 (1991); Zebedee, et al., Human Combinatorial Antibody Libraries to Hepatitis B Surface Antigen, *Proc. Natl. Acad. Sci. U.S.A.*, 89:3175-3179 (1992); Zuckerman, et al., Identification of Highest-Affinity Ligands by Affinity Selection from Equimolar Peptide Mixtures
- 15 Generated by Robotic Synthesis, *Proc. Natl. Acad. Sci. U.S.A.*, 89:4505-4509 (1992).

For example, peptides that bind to an MTSP9 polypeptide or a protease domain of an SP protein can be identified using phage display libraries. In an exemplary embodiment, this method can include a) contacting phage from a

20 phage library with the MTSP9 polypeptide or a protease domain thereof; (b) isolating phage that bind to the protein; and (c) determining the identity of at least one peptide coded by the isolated phage to identify a peptide that binds to an MTSP9 polypeptide.

H. Modulators of the activity of MTSP9 polypeptides

- 25 Provided herein are compounds, identified by screening or produced using the MTSP9 polypeptide or protease domain in other screening methods, that modulate the activity of an MTSP9. These compounds act by directly interacting with the MTSP9 polypeptide or by altering transcription or translation thereof. Such molecules include, but are not limited to, antibodies that specifically react
- 30 with an MTSP9 polypeptide, particularly with the protease domain thereof, antisense nucleic acids or double-stranded RNA (dsRNA) such as RNAi, that alter

-93-

expression of the MTSP9 polypeptide, antibodies, peptide mimetics and other such compounds.

1. Antibodies

Antibodies, including polyclonal and monoclonal antibodies, that
5 specifically bind to the MTSP9 polypeptide provided herein, particularly to the single chain protease domains thereof or the activated forms of the full-length or protease domain or the zymogen form, are provided.

Generally, the antibody is a monoclonal antibody, and typically the antibody specifically binds to the protease domain of the MTSP9 polypeptide. In
10 particular embodiments, antibodies to each of the single chain and or two chain form of the protease domain of MTSP9 are provided. Also provided are antibodies that specifically bind to any domain of MTSP9 and to two chain forms thereof.

The MTSP9 polypeptide and domains, fragments, homologs and
15 derivatives thereof can be used as immunogens to generate antibodies that specifically bind such immunogens. Such antibodies include but are not limited to polyclonal, monoclonal, chimeric, single chain, Fab fragments, and an Fab expression library. In a specific embodiment, antibodies to human MTSP9 polypeptide are produced. In another embodiment, complexes formed from
20 fragments of MTSP9 polypeptide, which fragments contain the serine protease domain, are used as immunogens for antibody production.

Various procedures known in the art can be used for the production of polyclonal antibodies to MTSP9 polypeptide, its domains, derivatives, fragments or analogs. For production of the antibody, various host animals can be
25 immunized by injection with the native MTSP9 polypeptide or a synthetic version, or a derivative of the foregoing, such as a cross-linked MTSP9 polypeptide. Such host animals include but are not limited to rabbits, mice, rats, etc. Various adjuvants can be used to increase the immunological response, depending on the host species, and include but are not limited to Freund's
30 (complete and incomplete), mineral gels such as aluminum hydroxide, surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil

-94-

emulsions, dinitrophenol, and potentially useful human adjuvants such as bacille Calmette-Guerin (BCG) and corynebacterium parvum.

For preparation of monoclonal antibodies directed towards an MTSP9 polypeptide or domains, derivatives, fragments or analogs thereof, any technique that provides for the production of antibody molecules by continuous cell lines in culture can be used. Such techniques include but are not restricted to the hybridoma technique originally developed by Kohler and Milstein (*Nature* 256:495-497 (1975)), the trioma technique, the human B-cell hybridoma technique (Kozbor et al., *Immunology Today* 4:72 (1983)), and the EBV hybridoma technique to produce human monoclonal antibodies (Cole et al., *in Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc., pp. 77-96 (1985)). In an additional embodiment, monoclonal antibodies can be produced in germ-free animals utilizing recent technology (PCT/US90/02545). Human antibodies can be used and can be obtained by using human hybridomas (Cote et al., *Proc. Natl. Acad. Sci. USA* 80:2026-2030 (1983)), or by transforming human B cells with EBV virus *in vitro* (Cole et al., *in Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc., pp. 77-96 (1985)). Techniques developed for the production of "chimeric antibodies" (Morrison et al., *Proc. Natl. Acad. Sci. USA* 81:6851-6855 (1984); Neuberger et al., *Nature* 312:604-608 (1984); Takeda et al., *Nature* 314:452-454 (1985)) by splicing the genes from a mouse antibody molecule specific for the MTSP9 polypeptide together with genes from a human antibody molecule of appropriate biological activity can be used.

MTSP9-encoding nucleic acid molecules or portions thereof can be used in DNA immunization protocols to produce antibodies that bind to MTSP9 (see, e.g., U.S. Patent No. 5,795,872 and U.S. Patent No. 5,643,578 and U.S. Patent No. 6,337,072).

Techniques described for the production of single chain antibodies (U.S. patent 4,946,778) can be adapted to produce MTSP9 polypeptide-specific single chain antibodies. An additional embodiment uses the techniques described for the construction of Fab expression libraries (Huse et al., *Science* 246:1275-1281 (1989)) to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity for MTSP9 polypeptide or domains, derivatives, or analogs

-95-

thereof. Non-human antibodies can be "humanized" by known methods (*see, e.g.,* U.S. Patent No. 5,225,539).

Antibody fragments that specifically bind to MTSP9 polypeptide or epitopes thereof can be generated by techniques known in the art. For example, such fragments include but are not limited to: the F(ab')₂ fragment, which can be produced by pepsin digestion of the antibody molecule; the Fab' fragments that can be generated by reducing the disulfide bridges of the F(ab')₂ fragment, the Fab fragments that can be generated by treating the antibody molecular with papain and a reducing agent, and Fv fragments.

In the production of antibodies, screening for the desired antibody can be accomplished by techniques known in the art, *e.g.,* ELISA (enzyme-linked immunosorbent assay). To select antibodies specific for a particular domain of the MTSP9 polypeptide one can assay generated hybridomas for a product that binds to the fragment of the MTSP9 polypeptide that contains such a domain.

The foregoing antibodies can be used in methods known in the art relating to the localization and/or quantitation of MTSP9 polypeptide proteins, *e.g.,* for imaging these proteins, measuring levels thereof in appropriate physiological samples, in, for example, diagnostic methods. In another embodiment, anti-MTSP9 polypeptide antibodies, or fragments thereof, containing the binding domain are used as therapeutic agents.

2. Peptides, Polypeptides and Peptide Mimetics

Provided herein are methods for identifying molecules that bind to and modulate the activity of SP proteins. Included among molecules that bind to SPs, particularly the single chain protease domain or catalytically active fragments thereof, are peptides, polypeptides and peptide mimetics, including cyclic peptides. Peptide mimetics are molecules or compounds that mimic the necessary molecular conformation of a ligand or polypeptide for specific binding to a target molecule such as an MTSP9 polypeptide. In an exemplary embodiment, the peptides, polypeptides and peptide mimetics or peptide mimetics bind to the protease domain of the MTSP9 polypeptide. Such peptides and peptide mimetics include those of antibodies that specifically bind to an MTSP9 polypeptide and, typically, bind to the protease domain of an

-96-

MTSP9 polypeptide. The peptides, polypeptides and peptide mimetics identified by methods provided herein can be agonists or antagonists of MTSP9 polypeptides.

Such peptides, polypeptides and peptide mimetics are useful for
5 diagnosing, treating, preventing, and screening for a disease or disorder associated with MTSP9 polypeptide activity in a mammal. In addition, the peptides and peptide mimetics are useful for identifying, isolating, and purifying molecules or compounds that modulate the activity of an MTSP9 polypeptide, or specifically bind to an MTSP9 polypeptide, generally the protease domain of an
10 MTSP9 polypeptide. Low molecular weight peptides and peptide mimetics can have strong binding properties to a target molecule, *e.g.*, an MTSP9 polypeptide or the protease domain of an MTSP9 polypeptide.

Peptides, polypeptides and peptide mimetics that bind to MTSP9 polypeptides as described herein can be administered to mammals, including
15 humans, to modulate MTSP9 polypeptide activity. Thus, methods for therapeutic treatment and prevention of neoplastic diseases comprise administering a peptide, polypeptide or peptide mimetic compound in an amount sufficient to modulate such activity are provided. Thus, also provided herein are methods for treating a subject having such a disease or disorder in which a peptide,
20 polypeptide or peptide mimetic compound is administered to the subject in a therapeutically effective dose or amount.

Compositions containing the peptides, polypeptides or peptide mimetics can be administered for prophylactic and/or therapeutic treatments. In therapeutic applications, compositions can be administered to a patient already
25 suffering from a disease, as described above, in an amount sufficient to cure or at least partially arrest the symptoms of the disease and its complications. Amounts effective for this use will depend on the severity of the disease and the weight and general state of the patient and can be empirically determined.

In prophylactic applications, compositions containing the peptides,
30 polypeptides and peptide mimetics are administered to a patient susceptible to or otherwise at risk of a particular disease. Such an amount is defined to be a "prophylactically effective dose". In this use, the precise amounts again depend

-97-

on the patient's state of health and weight. Accordingly, the peptides, polypeptides and peptide mimetics that bind to an MTSP9 polypeptide can be used to prepare pharmaceutical compositions containing, as an active ingredient, at least one of the peptides or peptide mimetics in association with a

5 pharmaceutical carrier or diluent. The compounds can be administered, for example, by oral, pulmonary, parental (intramuscular, intraperitoneal, intravenous (IV) or subcutaneous injection), inhalation (via a fine powder formulation), transdermal, nasal, vaginal, rectal, or sublingual routes of administration and can be formulated in dosage forms appropriate for each route of administration (see,

10 *e.g.*, International PCT application Nos. WO 93/25221 and WO 94/17784; and European Patent Application 613,683).

Peptides, polypeptides and peptide mimetics that bind to MTSP9 polypeptides are useful *in vitro* as unique tools for understanding the biological role of MTSP9 polypeptides, including the evaluation of the many factors

15 thought to influence, and be influenced by, the production of MTSP9 polypeptide. Such peptides, polypeptides and peptide mimetics are also useful in the development of other compounds that bind to and modulate the activity of an MTSP9 polypeptide, because such compounds provide important information on the relationship between structure and activity that should facilitate such

20 development.

The peptides, polypeptides and peptide mimetics are also useful as competitive binders in assays to screen for new MTSP9 polypeptides or MTSP9 polypeptide agonists. In such assay embodiments, the compounds can be used without modification or can be modified in a variety of ways; for example, by

25 labeling, such as covalently or non-covalently joining a moiety which directly or indirectly provides a detectable signal. In any of these assays, the materials thereto can be labeled either directly or indirectly. Possibilities for direct labeling include label groups such as: radiolabels such as ¹²⁵I enzymes (U.S. Pat. No. 3,645,090) such as peroxidase and alkaline phosphatase, and fluorescent labels

30 (U.S. Pat. No. 3,940,475) capable of monitoring the change in fluorescence intensity, wavelength shift, or fluorescence polarization. Possibilities for indirect labeling include biotinylation of one constituent followed by binding to avidin

-98-

coupled to one of the above label groups. The compounds can also include spacers or linkers in cases where the compounds are to be attached to a solid support.

Moreover, based on their ability to bind to an MTSP9 polypeptide, the peptides, polypeptides and peptide mimetics can be used as reagents for detecting MTSP9 polypeptides in living cells, fixed cells, in biological fluids, in tissue homogenates and in purified, natural biological materials. For example, by labelling such peptides, polypeptides and peptide mimetics, cells having MTSP9 polypeptides can be identified. In addition, based on their ability to bind an MTSP9 polypeptide, the peptides, polypeptides and peptide mimetics can be used in *in situ* staining, FACS (fluorescence-activated cell sorting), Western blotting, ELISA and other analytical protocols. Based on their ability to bind to an MTSP9 polypeptide, the peptides, polypeptides and peptide mimetics can be used in purification of MTSP9 polypeptides or in purifying cells expressing the MTSP9 polypeptides, *e.g.*, a polypeptide encoding the protease domain of an MTSP9 polypeptide.

The peptides, polypeptides and peptide mimetics can also be used as commercial reagents for various medical research and diagnostic uses. The activity of the peptides and peptide mimetics can be evaluated either *in vitro* or *in vivo* in one of the numerous models described in McDonald (1992) *Am. J. of Pediatric Hematology/Oncology*, 14:8-21.

3. Peptide, polypeptide and peptide mimetic therapy

Peptide analogs are commonly used in the pharmaceutical industry as non-peptide drugs with properties analogous to those of the template peptide. These types of non-peptide compounds are termed "peptide mimetics" or "peptidomimetics" (Luthman *et al.*, *A Textbook of Drug Design and Development*, 14:386-406, 2nd Ed., Harwood Academic Publishers (1996); Joachim Grante (1994) *Angew. Chem. Int. Ed. Engl.*, 33:1699-1720; Fauchere (1986) *J. Adv. Drug Res.*, 15:29; Veber and Freidinger (1985) *TINS*, p. 392; and Evans *et al.* (1987) *J. Med. Chem.* 30:1229). Peptide mimetics that are structurally similar to therapeutically useful peptides can be used to produce an

-99-

equivalent or enhanced therapeutic or prophylactic effect. Preparation of peptidomimetics and structures thereof are known to those of skill in this art.

Systematic substitution of one or more amino acids of a consensus sequence with a D-amino acid of the same type (*e.g.*, D-lysine in place of
5 L-lysine) can be used to generate more stable peptides. In addition, constrained peptides containing a consensus sequence or a substantially identical consensus sequence variation can be generated by methods known in the art (Rizo *et al.* (1992) *An. Rev. Biochem.*, 61:387, incorporated herein by reference); for example, by adding internal cysteine residues capable of forming intramolecular
10 disulfide bridges which cyclize the peptide.

Those skilled in the art appreciate that modifications can be made to the peptides and mimetics without deleteriously effecting the biological or functional activity of the peptide. Further, the skilled artisan would know how to design non-peptide structures in three dimensional terms, that mimic the peptides that
15 bind to a target molecule, *e.g.*, an MTSP9 polypeptide or, generally, the protease domain of MTSP9 polypeptides (see, *e.g.*, Eck and Sprang (1989) *J. Biol. Chem.*, 26: 17605-18795).

When used for diagnostic purposes, the peptides and peptide mimetics can be labeled with a detectable label and, accordingly, the peptides and peptide
20 mimetics without such a label can serve as intermediates in the preparation of labeled peptides and peptide mimetics. Detectable labels can be molecules or compounds, which when covalently attached to the peptides and peptide mimetics, permit detection of the peptide and peptide mimetics *in vivo*, for example, in a patient to whom the peptide or peptide mimetic has been
25 administered, or *in vitro*, *e.g.*, in a sample or cells. Suitable detectable labels are well known in the art and include, by way of example, radioisotopes, fluorescent labels (*e.g.*, fluorescein), and the like. The particular detectable label employed is not critical and is selected to be detectable at non-toxic levels. Selection of the such labels is well within the skill of the art.

30 Covalent attachment of a detectable label to the peptide or peptide mimetic is accomplished by conventional methods well known in the art. For example, when the ¹²⁵I radioisotope is employed as the detectable label, covalent

-100-

attachment of ^{125}I to the peptide or the peptide mimetic can be achieved by incorporating the amino acid tyrosine into the peptide or peptide mimetic and then iodinating the peptide (see, *e.g.*, Weaner *et al.* (1994) *Synthesis and Applications of Isotopically Labelled Compounds*, pp. 137-140). If tyrosine is
5 not present in the peptide or peptide mimetic, incorporation of tyrosine to the N or C terminus of the peptide or peptide mimetic can be achieved by well known chemistry. Likewise, ^{32}P can be incorporated onto the peptide or peptide mimetic as a phosphate moiety through, for example, a hydroxyl group on the peptide or peptide mimetic using conventional chemistry.

10 Labeling of peptidomimetics usually involves covalent attachment of one or more labels, directly or through a spacer (*e.g.*, an amide group), to non-interfering position(s) on the peptidomimetic that are predicted by quantitative structure-activity data and/or molecular modeling. Such non-interfering positions generally are positions that do not form direct contacts
15 with the macromolecules(s) to which the peptidomimetic binds to produce the therapeutic effect. Derivatization (*e.g.*, labeling) of peptidomimetics should not substantially interfere with the desired biological or pharmacological activity of the peptidomimetic.

Peptides, polypeptides and peptide mimetics that can bind to an MTSP9
20 polypeptide or the protease domain of MTSP9 polypeptides and/or modulate the activity thereof, or exhibit MTSP9 polypeptide activity, can be used for treatment of neoplastic disease. The peptides, polypeptides and peptide mimetics can be delivered, *in vivo* or *ex vivo*, to the cells of a subject in need of treatment. Further, peptides which have MTSP9 polypeptide activity can be delivered, *in*
25 *vivo* or *ex vivo*, to cells which carry mutant or missing alleles encoding the MTSP9 polypeptide gene. Any of the techniques described herein or known to the skilled artisan can be used for preparation and *in vivo* or *ex vivo* delivery of such peptides, polypeptides and peptide mimetics that are substantially free of other human proteins. For example, the peptides, polypeptides and peptide
30 mimetics can be readily prepared by expression in a microorganism or synthesis *in vitro*.

-101-

The peptides or peptide mimetics can be introduced into cells, *in vivo* or *ex vivo*, by microinjection or by use of liposomes, for example. Alternatively, the peptides, polypeptides or peptide mimetics can be taken up by cells, *in vivo* or *ex vivo*, actively or by diffusion. In addition, extracellular application of the peptide, polypeptide or peptide mimetic can be sufficient to effect treatment of a neoplastic disease. Other molecules, such as drugs or organic compounds, that: 1) bind to a MTSP9 polypeptide or protease domain thereof; or 2) have a similar function or activity to an MTSP9 polypeptide or protease domain thereof, can be used in methods for treatment.

4. Rational drug design

The goal of rational drug design is to produce structural analogs of biologically active polypeptides or peptides of interest or of small molecules or peptide mimetics with which they interact (*e.g.*, agonists and antagonists) in order to fashion drugs which are, *e.g.*, more active or stable forms thereof; or which, for example, enhance or interfere with the function of a polypeptide *in vivo* (*e.g.*, an MTSP9 polypeptide). In one approach, one first determines the three-dimensional structure of a protein of interest (*e.g.*, an MTSP9 polypeptide or polypeptide having a protease domain) or, for example, of an MTSP9 polypeptide-ligand complex, by X-ray crystallography, by computer modeling or most typically, by a combination of approaches (see, *e.g.*, Erickson *et al.* 1990). Also, useful information regarding the structure of a polypeptide can be gained by modeling based on the structure of homologous proteins. In addition, peptides can be analyzed by an alanine scan. In this technique, an amino acid residue is replaced by Ala, and its effect on the peptide's activity is determined. Each of the amino acid residues of the peptide is analyzed in this manner to determine the important regions of the peptide.

Also, a polypeptide or peptide that binds to an MTSP9 polypeptide or, generally, the protease domain of an MTSP9 polypeptide, can be selected by a functional assay, and then the crystal structure of this polypeptide or peptide can be determined. The polypeptide can be, for example, an antibody specific for an MTSP9 polypeptide or the protein domain of an MTSP9 polypeptide. This approach can yield a pharmacophore upon which subsequent drug design can be

-102-

based. Further, it is possible to bypass the crystallography altogether by generating anti-idiotypic polypeptides or peptides, (anti-ids) to a functional, pharmacologically active polypeptide or peptide that binds to an MTSP9 polypeptide or protease domain of an MTSP9 polypeptide. As a mirror image of a mirror image, the binding site of the anti-ids is expected to be an analog of the original target molecule, e.g., an MTSP9 polypeptide or polypeptide having an MTSP9 polypeptide. The anti-id could then be used to identify and isolate peptides from banks of chemically or biologically produced banks of peptides. Selected peptides would then act as the pharmacophore.

Thus, one can design drugs which have, for example, improved activity or stability or which act as modulators (*e.g.*, inhibitors, agonists or antagonists) of MTSP9 polypeptide activity, and are useful in the methods, particularly the methods for diagnosis, treatment, prevention, and screening of a neoplastic disease. By virtue of the availability of nucleic acid that encodes MTSP9 polypeptides, sufficient amounts of the MTSP9 polypeptide can be made available to perform such analytical studies as X-ray crystallography. In addition, the knowledge of the amino acid sequence of an MTSP9 polypeptide or the protease domain thereof, *e.g.*, the protease domain encoded by the amino acid sequence of SEQ ID Nos. 5 and 6, can provide guidance on computer modeling techniques in place of, or in addition to, X-ray crystallography.

Methods of identifying peptides and peptide mimetics that bind to MTSP9 polypeptides

Peptides having a binding affinity to the MTSP9 polypeptide provided herein (*e.g.*, an MTSP9 polypeptide or a polypeptide having a protease domain of an MTSP9 polypeptide) can be readily identified, for example, by random peptide diversity generating systems coupled with an affinity enrichment process. Specifically, random peptide diversity generating systems include the "peptides on plasmids" system (see, *e.g.*, U.S. Patent Nos. 5,270,170 and 5,338,665); the "peptides on phage" system (see, *e.g.*, U.S. Patent No. 6,121,238 and Cwirla, *et al.* (1990) *Proc. Natl. Acad. Sci. U.S.A.* 87:6378-6382); the "polysome system;" the "encoded synthetic library (ESL)" system; and the "very

-103-

large scale immobilized polymer synthesis" system (see, *e.g.*, U.S. Patent No. 6,121,238; and Dower *et al.* (1991) *An. Rep. Med. Chem.* 26:271-280

For example, using the procedures described above, random peptides can generally be designed to have a defined number of amino acid residues in length
5 (*e.g.*, 12). To generate the collection of oligonucleotides encoding the random peptides, the codon motif (NNK)_x, where N is nucleotide A, C, G, or T (equimolar; depending on the methodology employed, other nucleotides can be employed), K is G or T (equimolar), and x is an integer corresponding to the number of amino acids in the peptide (*e.g.*, 12) can be used to specify any one
10 of the 32 possible codons resulting from the NNK motif: 1 for each of 12 amino acids, 2 for each of 5 amino acids, 3 for each of 3 amino acids, and only one of the three stop codons. Thus, the NNK motif encodes all of the amino acids, encodes only one stop codon, and reduces codon bias.

The random peptides can be presented, for example, either on the surface
15 of a phage particle, as part of a fusion protein containing either the pIII or the pVIII coat protein of a phage fd derivative (peptides on phage) or as a fusion protein with the LacI peptide fusion protein bound to a plasmid (peptides on plasmids). The phage or plasmids, including the DNA encoding the peptides, can be identified and isolated by an affinity enrichment process using immobilized
20 MTSP9 polypeptide having a protease domain. The affinity enrichment process, sometimes called "panning," typically involves multiple rounds of incubating the phage, plasmids, or polysomes with the immobilized MTSP9 polypeptide, collecting the phage, plasmids, or polysomes that bind to the MTSP9 polypeptide (along with the accompanying DNA or mRNA), and producing more of the phage
25 or plasmids (along with the accompanying LacI-peptide fusion protein) collected.

Characteristics of peptides and peptide mimetics

Among the peptides, polypeptides and peptide mimetics for therapeutic application are those of having molecular weights from about 250 to about 8,000 daltons. If such peptides are oligomerized, dimerized and/or derivatized
30 with a hydrophilic polymer (*e.g.*, to increase the affinity and/or activity of the compounds), the molecular weights of such peptides can be substantially greater and can range anywhere from about 500 to about 120,000 daltons, generally

-104-

from about 8,000 to about 80,000 daltons. Such peptides can contain 9 or more amino acids that are naturally occurring or synthetic (non-naturally occurring) amino acids. One skilled in the art can determine the affinity and molecular weight of the peptides and peptide mimetics suitable for therapeutic and/or diagnostic purposes (*e.g.*, see Dower *et al.*, U.S. Patent No. 6,121,238).

The peptides can be covalently attached to one or more of a variety of hydrophilic polymers. Suitable hydrophilic polymers include, but are not limited to, polyalkylethers as exemplified by polyethylene glycol and polypropylene glycol, polylactic acid, polyglycolic acid, polyoxyalkenes, polyvinylalcohol, polyvinylpyrrolidone, cellulose and cellulose derivatives, dextran and dextran derivatives. When the peptide compounds are derivatized with such polymers, their solubility and circulation half-lives can be increased with little, if any, diminishment in their binding activity. The peptide compounds can be dimerized and each of the dimeric subunits can be covalently attached to a hydrophilic polymer. The peptide compounds can be PEGylated, *i.e.*, covalently attached to polyethylene glycol (PEG).

5. Methods of preparing peptides and peptide mimetics

Peptides that bind to MTSP9 polypeptides can be prepared by classical methods known in the art, for example, by using standard solid phase techniques. The standard methods include exclusive solid phase synthesis, partial solid phase synthesis methods, fragment condensation, classical solution synthesis, and even by recombinant DNA technology (see, *e.g.*, Merrifield (1963) *J. Am. Chem. Soc.*, 85:2149, incorporated herein by reference.)

Using the "encoded synthetic library" or "very large scale immobilized polymer synthesis" systems (see, *e.g.*, U.S. Patent No. 5,925,525, and 5,902,723), the minimum size of a peptide with the activity of interest can be determined. In addition all peptides that form the group of peptides that differ from the desired motif (or the minimum size of that motif) in one, two, or more residues can be prepared. This collection of peptides then can be screened for the ability to bind to the target molecule, *e.g.*, MTSP9 polypeptide or, generally, the protease domain of an MTSP9 polypeptide. This immobilized polymer synthesis system or other peptide synthesis methods can also be used to

-105-

synthesize truncation analogs and deletion analogs and combinations of truncation and deletion analogs of the peptide compounds.

These procedures can also be used to synthesize peptides in which amino acids other than the 20 naturally occurring, genetically encoded amino acids are substituted at one, two, or more positions of the peptide. For instance, naphthylalanine can be substituted for tryptophan, facilitating synthesis. Other synthetic amino acids that can be substituted into the peptides include L-hydroxypropyl, L-3, 4-dihydroxy-phenylalanyl, d amino acids such as L-d-hydroxylysyl and D-d-methylalanyl, L- α -methylalanyl, β amino acids, and isoquinolyl. D amino acids and non-naturally occurring synthetic amino acids can also be incorporated into the peptides (see, *e.g.*, Roberts *et al.* (1983) *Unusual Amino/Acids in Peptide Synthesis*, 5(6):341-449).

The peptides also can be modified by phosphorylation (see, *e.g.*, W. Bannwarth *et al.* (1996) *Biorganic and Medicinal Chemistry Letters*, 6(17):2141-2146), and other methods for making peptide derivatives (see, *e.g.*, Hruby *et al.* (1990) *Biochem. J.*, 268(2):249-262). Thus, peptide compounds also serve as a basis to prepare peptide mimetics with similar or improved biological activity.

Those of skill in the art recognize that a variety of techniques are available for constructing peptide mimetics with the same or similar desired biological activity as the corresponding peptide compound but with more favorable activity than the peptide with respect to solubility, stability, and susceptibility to hydrolysis and proteolysis (see, *e.g.*, Morgan *et al.* (1989) *An. Rep. Med. Chem.*, 24:243-252). Methods for preparing peptide mimetics modified at the N-terminal amino group, the C-terminal carboxyl group, and/or changing one or more of the amido linkages in the peptide to a non-amido linkage are known to those of skill in the art.

Amino terminus modifications include, but are not limited to, alkylating, acetylating and adding a carbobenzoyl group, forming a succinimide group (see, *e.g.*, Murray *et al.* (1995) *Burger's Medicinal Chemistry and Drug Discovery*, 5th ed., Vol. 1, Manfred E. Wolf, ed., John Wiley and Sons, Inc.). C-terminal

-106-

modifications include mimetics wherein the C-terminal carboxyl group is replaced by an ester, an amide or modifications to form a cyclic peptide.

In addition to N-terminal and C-terminal modifications, the peptide compounds, including peptide mimetics, can advantageously be modified with or covalently coupled to one or more of a variety of hydrophilic polymers. It has been found that when peptide compounds are derivatized with a hydrophilic polymer, their solubility and circulation half-lives can be increased and their immunogenicity is masked, with little, if any, diminishment in their binding activity. Suitable nonproteinaceous polymers include, but are not limited to, polyalkylethers as exemplified by polyethylene glycol and polypropylene glycol, polylactic acid, polyglycolic acid, polyoxyalkenes, polyvinylalcohol, polyvinylpyrrolidone, cellulose and cellulose derivatives, dextran and dextran derivatives. Generally, such hydrophilic polymers have an average molecular weight ranging from about 500 to about 100,000 daltons, including from about 2,000 to about 40,000 daltons and, from about 5,000 to about 20,000 daltons. The hydrophilic polymers also can have an average molecular weights of about 5,000 daltons, 10,000 daltons and 20,000 daltons.

Methods for derivatizing peptide compounds or for coupling peptides to such polymers have been described (see, *e.g.*, Zallipsky (1995) *Bioconjugate Chem.*, 6:150-165; Monfardini *et al.* (1995) *Bioconjugate Chem.*, 6:62-69; U.S. Pat. No. 4,640,835; U.S. Pat. No. 4,496,689; U.S. Pat. No. 4,301,144; U.S. Pat. No. 4,670,417; U.S. Pat. No. 4,791,192; U.S. Pat. No. 4,179,337 and WO 95/34326, all of which are incorporated by reference in their entirety herein).

Other methods for making peptide derivatives are described, for example, in Hruby *et al.* (1990), *Biochem J.*, 268(2):249-262, which is incorporated herein by reference. Thus, the peptide compounds also serve as structural models for non-peptidic compounds with similar biological activity. Those of skill in the art recognize that a variety of techniques are available for constructing compounds with the same or similar desired biological activity as a particular peptide compound but with more favorable activity with respect to solubility, stability, and susceptibility to hydrolysis and proteolysis (see, *e.g.*, Morgan *et al.* (1989) *An. Rep. Med. Chem.*, 24:243-252, incorporated herein by

-107-

reference). These techniques include replacing the peptide backbone with a backbone composed of phosphonates, amidates, carbamates, sulfonamides, secondary amines, and N-methylamino acids.

Peptide compounds can exist in a cyclized form with an intramolecular
5 disulfide bond between the thiol groups of the cysteines. Alternatively, an intermolecular disulfide bond between the thiol groups of the cysteines can be produced to yield a dimeric (or higher oligomeric) compound. One or more of the cysteine residues can also be substituted with a homocysteine.

I. Conjugates

10 A conjugate, containing: a) a single chain protease domain (or proteolytically active portion thereof) of an MTSP9 polypeptide or a full length zymogen, activated form thereof, or two or single chain protease domain thereof; and b) a targeting agent linked to the MTSP9 polypeptide directly or via a linker, wherein the agent facilitates: i) affinity isolation or purification of the
15 conjugate; ii) attachment of the conjugate to a surface; iii) detection of the conjugate; or iv) targeted delivery to a selected tissue or cell, is provided herein. The conjugate can be a chemical conjugate or a fusion protein mixture thereof.

The targeting agent can be a protein or peptide fragment, such as a tissue specific or tumor specific monoclonal antibody or growth factor or
20 fragment thereof linked either directly or via a linker to an MTSP9 polypeptide or a protease domain thereof. The targeting agent can also be a protein or peptide fragment that contains a protein binding sequence, a nucleic acid binding sequence, a lipid binding sequence, a polysaccharide binding sequence, or a metal binding sequence, or a linker for attachment to a solid support. In a
25 particular embodiment, the conjugate contains a) the MTSP9 or portion thereof, as described herein; and b) a targeting agent linked to the MTSP9 polypeptide directly or via a linker.

Conjugates, such as fusion proteins and chemical conjugates, of the MTSP9 polypeptide with a protein or peptide fragment (or plurality thereof) that
30 functions, for example, to facilitate affinity isolation or purification of the MTSP9 polypeptide domain, attachment of the MTSP9 polypeptide domain to a surface, or detection of the MTSP9 polypeptide domain are provided. The conjugates can

-108-

be produced by chemical conjugation, such as via thiol linkages, and can be produced by recombinant means as fusion proteins. In the fusion protein, the peptide or fragment thereof is linked to either the N-terminus or C-terminus of the MTSP9 polypeptide domain. In chemical conjugates the peptide or fragment
5 thereof can be linked anywhere that conjugation can be effected, and there can be a plurality of such peptides or fragments linked to a single MTSP9 polypeptide domain or to a plurality thereof.

The targeting agent is for *in vitro* or *in vivo* delivery to a cell or tissue, and includes agents such as cell or tissue-specific antibodies, growth factors and
10 other factors that bind to moieties expressed on specific cells; and other cell or tissue specific agents that promote directed delivery of a linked protein. The targeting agent can be one that specifically delivers the MTSP9 polypeptide to selected cells by interaction with a cell surface protein and internalization of conjugate or MTSP9 polypeptide portion thereof.

15 These conjugates are used in a variety of methods and are particularly suited for use in methods of activation of prodrugs, such as prodrugs that upon cleavage by the particular MTSP9, which is localized at or near the targeted cell or tissue, are cytotoxic. The prodrugs are administered prior to, or simultaneously with, or subsequently to the conjugate. Upon delivery to the
20 targeted cells, the protease activates the prodrug, which then exhibits a therapeutic effect, such as a cytotoxic effect.

1. Conjugation

Conjugates with linked MTSP9 polypeptide domains can be prepared either by chemical conjugation, recombinant DNA technology, or combinations of
25 recombinant expression and chemical conjugation. The MTSP9 polypeptide domains and the targeting agent can be linked in any orientation and more than one targeting agents and/or MTSP9 polypeptide domains can be present in a conjugate.

-109-

a. Fusion proteins

Fusion proteins are provided herein. A fusion protein contains: a) one or a plurality of domains of an MTSP9 polypeptide and b) a targeting agent. The fusion proteins are generally produced by recombinant expression of nucleic acids that encode the fusion protein.

b. Chemical conjugation

To effect chemical conjugation herein, the MTSP9 polypeptide domain is linked via one or more selected linkers or directly to the targeting agent. Chemical conjugation must be used if the targeted agent is other than a peptide or protein, such as a nucleic acid or a non-peptide drug. Any means known to those of skill in the art for chemically conjugating selected moieties can be used.

2. Linkers

Linkers for two purposes are contemplated herein. The conjugates can include one or more linkers between the MTSP9 polypeptide portion and the targeting agent. Additionally, linkers are used for facilitating or enhancing immobilization of an MTSP9 polypeptide or portion thereof on a solid support, such as a microtiter plate, silicon or silicon-coated chip, glass or plastic support, such as for high throughput solid phase screening protocols.

Any linker known to those of skill in the art for preparation of conjugates can be used herein. These linkers are typically used in the preparation of chemical conjugates; peptide linkers can be incorporated into fusion proteins.

Linkers can be any moiety suitable to associate a domain of MTSP9 polypeptide and a targeting agent. Such linkers and linkages include, but are not limited to, peptidic linkages, amino acid and peptide linkages, typically containing between one and about 60 amino acids, more generally between about 10 and 30 amino acids, chemical linkers, such as heterobifunctional cleavable cross-linkers, including but are not limited to, N-succinimidyl (4-iodoacetyl)-aminobenzoate, sulfosuccinimidyl (4-iodoacetyl)-aminobenzoate, 4-succinimidyl-oxycarbonyl-a- (2-pyridyldithio)toluene, sulfosuccinimidyl-6- [a-methyl-a- (pyridyldithiol)-toluamido] hexanoate, N-succinimidyl-3-(-2-pyridyldithio) -propionate, succinimidyl 6[3-(-2-pyridyldithio)-propionamido] hexanoate, sulfosuccinimidyl 6[3-(-2-pyridyldithio)-propionamido] hexanoate, 3-(2-pyridyldi-

-110-

thio)-propionyl hydrazide, Ellman's reagent, dichlorotriazinic acid, and S-(2-thiopyridyl)-L-cysteine. Other linkers include, but are not limited to peptides and other moieties that reduce steric hindrance between the domain of MTSP9 polypeptide and the targeting agent, intracellular enzyme substrates, linkers that
5 increase the flexibility of the conjugate, linkers that increase the solubility of the conjugate, linkers that increase the serum stability of the conjugate, photocleavable linkers and acid cleavable linkers.

Other exemplary linkers and linkages that are suitable for chemically linked conjugates include, but are not limited to, disulfide bonds, thioether
10 bonds, hindered disulfide bonds, and covalent bonds between free reactive groups, such as amine and thiol groups. These bonds are produced using heterobifunctional reagents to produce reactive thiol groups on one or both of the polypeptides and then reacting the thiol groups on one polypeptide with reactive thiol groups or amine groups to which reactive maleimido groups or thiol
15 groups can be attached on the other. Other linkers include, acid cleavable linkers, such as bismaleimideoxy propane, acid labile-transferrin conjugates and adipic acid dihydrazide, that would be cleaved in more acidic intracellular compartments; cross linkers that are cleaved upon exposure to UV or visible light and linkers, such as the various domains, such as C_H1, C_H2, and C_H3, from
20 the constant region of human IgG₁ (see, Batra *et al. Molecular Immunol.*, 30:379-386 (1993)). In some embodiments, several linkers can be included in order to take advantage of desired properties of each linker.

Chemical linkers and peptide linkers can be inserted by covalently coupling the linker to the domain of MTSP9 polypeptide and the targeting agent.
25 The heterobifunctional agents, described below, can be used to effect such covalent coupling. Peptide linkers can also be linked by expressing DNA encoding the linker and therapeutic agent (TA), linker and targeted agent, or linker, targeted agent and therapeutic agent (TA) as a fusion protein. Flexible linkers and linkers that increase solubility of the conjugates are contemplated for
30 use, either alone or with other linkers are also contemplated herein.

-111-

a) Acid cleavable, photocleavable and heat sensitive linkers

Acid cleavable linkers, photocleavable and heat sensitive linkers can also be used, particularly where it can be necessary to cleave the domain of MTSP9 polypeptide to permit it to be more readily accessible to reaction. Acid cleavable
5 linkers include, but are not limited to, bismaleimideoxy propane; and adipic acid dihydrazide linkers (see, *e.g.*, Fattom *et al.* (1992) *Infection & Immun.* 60:584-589) and acid labile transferrin conjugates that contain a sufficient portion of transferrin to permit entry into the intracellular transferrin cycling pathway (see, *e.g.*, Welhöner *et al.* (1991) *J. Biol. Chem.* 266:4309-4314).

10 Photocleavable linkers are linkers that are cleaved upon exposure to light (see, *e.g.*, Goldmacher *et al.* (1992) *Bioconj. Chem.* 3:104-107, which linkers are herein incorporated by reference), thereby releasing the targeted agent upon exposure to light. Photocleavable linkers that are cleaved upon exposure to light are known (see, *e.g.*, Hazum *et al.* (1981) in *Pept., Proc. Eur. Pept. Symp.*,
15 16th, Brunfeldt, K (Ed), pp. 105-110, which describes the use of a nitrobenzyl group as a photocleavable protective group for cysteine; Yen *et al.* (1989) *Makromol. Chem* 190:69-82, which describes water soluble photocleavable copolymers, including hydroxypropylmethacrylamide copolymer, glycine copolymer, fluorescein copolymer and methylrhodamine copolymer; Gold-
20 macher *et al.* (1992) *Bioconj. Chem.* 3:104-107, which describes a cross-linker and reagent that undergoes photolytic degradation upon exposure to near UV light (350 nm); and Senter *et al.* (1985) *Photochem. Photobiol* 42:231-237, which describes nitrobenzyloxycarbonyl chloride cross linking reagents that produce photocleavable linkages), thereby releasing the targeted agent upon
25 exposure to light. Such linkers would have particular use in treating dermatological or ophthalmic conditions that can be exposed to light using fiber optics. After administration of the conjugate, the eye or skin or other body part can be exposed to light, resulting in release of the targeted moiety from the conjugate. Such photocleavable linkers are useful in connection with diagnostic

-112-

protocols in which it is desirable to remove the targeting agent to permit rapid clearance from the body of the animal.

b) Other linkers for chemical conjugation

Other linkers, include trityl linkers, particularly, derivatized
5 trityl groups to generate a genus of conjugates that provide for
release of therapeutic agents at various degrees of acidity or alkalinity.
The flexibility thus afforded by the ability to preselect the pH range at
which the therapeutic agent is released allows selection of a linker based on the
known physiological differences between tissues in need of delivery of a
10 therapeutic agent (see, *e.g.*, U.S. Patent No. 5,612,474). For example, the
acidity of tumor tissues appears to be lower than that of normal tissues.

c) Peptide linkers

The linker moieties can be peptides. Peptide linkers can be employed in
fusion proteins and also in chemically linked conjugates. The peptide typically
15 has from about 2 to about 60 amino acid residues, for example from about 5 to
about 40, or from about 10 to about 30 amino acid residues. The length
selected depends upon factors, such as the use for which the linker is included.

Peptide linkers are advantageous when the targeting agent is
proteinaceous. For example, the linker moiety can be a flexible spacer amino
20 acid sequence, such as those known in single-chain antibody research.
Examples of such known linker moieties include, but are not limited to,
peptides, such as $(\text{Gly}_m\text{Ser})_n$ and $(\text{Ser}_m\text{Gly})_n$, in which n is 1 to 6, including 1 to
4 and 2 to 4, and m is 1 to 6, including 1 to 4, and 2 to 4, enzyme cleavable
linkers and others.

25 Additional linking moieties are described, for example, in Huston *et al.*,
Proc. Natl. Acad. Sci. U.S.A. 85:5879-5883, 1988; Whitlow, M., *et al.*, *Protein*
Engineering 6:989-995, 1993; Newton *et al.*, *Biochemistry* 35:545-553, 1996;
A. J. Cumber *et al.*, *Bioconj. Chem.* 3:397-401, 1992; Ladurner *et al.*, *J. Mol.*
Biol. 273:330-337, 1997; and U.S. Patent. No. 4,894,443. In some
30 embodiments, several linkers can be included in order to take advantage of
desired properties of each linker.

-113-

3. Targeting agents

Any agent that facilitates detection, immobilization, or purification of the conjugate is contemplated for use herein. For chemical conjugates any moiety that has such properties is contemplated; for fusion proteins, the targeting agent is a protein, peptide or fragment thereof that is sufficient to effects the targeting activity. Contemplated targeting agents include those that deliver the MTSP9 polypeptide or portion thereof to selected cells and tissues. Such agents include tumor specific monoclonal antibodies and portions thereof, growth factors, such as FGF, EGF, PDGF, VEGF, cytokines, including chemokines, and other such agents.

4. Nucleic acids, plasmids and cells

Isolated nucleic acid fragments encoding fusion proteins are provided. The nucleic acid fragment that encodes the fusion protein includes: a) nucleic acid encoding a protease domain of an MTSP9 polypeptide; and b) nucleic acid encoding a protein, peptide or effective fragment thereof that facilitates: i) affinity isolation or purification of the fusion protein; ii) attachment of the fusion protein to a surface; or iii) detection of the fusion protein. Generally, the nucleic acid is DNA.

Plasmids for replication and vectors for expression that contain the above nucleic acid fragments are also provided. Cells containing the plasmids and vectors are also provided. The cells can be any suitable host including, but are not limited to, bacterial cells, yeast cells, fungal cells, plant cells, insect cell and animal cells. The nucleic acids, plasmids, and cells containing the plasmids can be prepared according to methods known in the art including any described herein.

Also provided are methods for producing the above fusion proteins. An exemplary method includes the steps of growing, for example, culturing the cells so that they proliferate, cells containing a plasmid encoding the fusion protein under conditions whereby the fusion protein is expressed by the cell, and recovering the expressed fusion protein. Methods for expressing and recovering recombinant proteins are well known in the art (*See generally, Current Protocols*

-114-

in Molecular Biology (1998) § 16, John Wiley & Sons, Inc.) and such methods can be used for expressing and recovering the expressed fusion proteins.

The recovered fusion proteins can be isolated or purified by methods known in the art such as centrifugation, filtration, chromatography, electrophoresis, immunoprecipitation, and other such methods, or by a combination thereof (*See generally, Current Protocols in Molecular Biology* (1998) § 10, John Wiley & Sons, Inc.). Generally the recovered fusion protein is isolated or purified through affinity binding between the protein or peptide fragment of the fusion protein and an affinity binding moiety. As discussed in the above sections regarding the construction of the fusion proteins, any affinity binding pairs can be constructed and used in the isolation or purification of the fusion proteins. For example, the affinity binding pairs can be protein binding sequences/protein, DNA binding sequences/DNA sequences, RNA binding sequences/RNA sequences, lipid binding sequences/lipid, polysaccharide binding sequences/polysaccharide, or metal binding sequences/metal.

5. Immobilization and supports or substrates therefor

In certain embodiments, where the targeting agents are designed for linkage to surfaces, the MTSP9 polypeptide can be attached by linkage such as ionic or covalent, non-covalent or other chemical interaction, to a surface of a support or matrix material. Immobilization can be effected directly or via a linker. The MTSP9 polypeptide can be immobilized on any suitable support, including, but are not limited to, silicon chips, and other supports described herein and known to those of skill in the art. A plurality of MTSP9 polypeptide or protease domains thereof can be attached to a support, such as an array (*i.e.*, a pattern of two or more) of conjugates on the surface of a silicon chip or other chip for use in high throughput protocols and formats.

It is also noted that the domains of the MTSP9 polypeptide can be linked directly to the surface or via a linker without a targeting agent linked thereto. Hence chips containing arrays of the domains of the MTSP9 polypeptide.

The matrix material or solid supports contemplated herein are generally any of the insoluble materials known to those of skill in the art to immobilize ligands and other molecules, and are those that used in many chemical

-115-

syntheses and separations. Such supports are used, for example, in affinity chromatography, in the immobilization of biologically active materials, and during chemical syntheses of biomolecules, including proteins, amino acids and other organic molecules and polymers. The preparation of and use of supports is well known to those of skill in this art; there are many such materials and preparations thereof known. For example, naturally-occurring support materials, such as agarose and cellulose, can be isolated from their respective sources, and processed according to known protocols, and synthetic materials can be prepared in accord with known protocols.

The supports are typically insoluble materials that are solid, porous, deformable, or hard, and have any required structure and geometry, including, but not limited to: beads, pellets, disks, capillaries, hollow fibers, needles, solid fibers, random shapes, thin films and membranes. Thus, the item can be fabricated from the matrix material or combined with it, such as by coating all or part of the surface or impregnating particles.

Typically, when the matrix is particulate, the particles are at least about 10-2000 μm , but can be smaller or larger, depending upon the selected application. Selection of the matrices is governed, at least in part, by their physical and chemical properties, such as solubility, functional groups, mechanical stability, surface area swelling propensity, hydrophobic or hydrophilic properties and intended use.

If necessary, the support matrix material can be treated to contain an appropriate reactive moiety. In some cases, the support matrix material already containing the reactive moiety can be obtained commercially. The support matrix material containing the reactive moiety can thereby serve as the matrix support upon which molecules are linked. Materials containing reactive surface moieties such as amino silane linkages, hydroxyl linkages or carboxysilane linkages can be produced by well established surface chemistry techniques involving silanization reactions, or the like. Examples of these materials are those having surface silicon oxide moieties, covalently linked to gamma-amino-propylsilane, and other organic moieties; N-[3-(triethoxysilyl)propyl]phthelamic acid; and bis-(2-hydroxyethyl)aminopropyltriethoxysilane. Exemplary of readily

-116-

available materials containing amino group reactive functionalities, include, but are not limited to, para-aminophenyltriethoxysilane. Also derivatized polystyrenes and other such polymers are well known and readily available to those of skill in this art (*e.g.*, the Tentagel® Resins are available with a multitude of functional groups, and are sold by Rapp Polymere, Tübingen, Germany; see, 5 U.S. Patent No. 4,908,405 and U.S. Patent No. 5,292,814; *see, also* Butz et al., *Peptide Res.*, 7:20-23 (1994); and Kleine et al., *Immunobiol.*, 190:53-66 (1994)).

These matrix materials include any material that can act as a support 10 matrix for attachment of the molecules of interest. Such materials are known to those of skill in this art, and include those that are used as a support matrix. These materials include, but are not limited to, inorganics, natural polymers, and synthetic polymers, including, but are not limited to: cellulose, cellulose derivatives, acrylic resins, glass, silica gels, polystyrene, gelatin, polyvinyl 15 pyrrolidone, co-polymers of vinyl and acrylamide, polystyrene cross-linked with divinylbenzene and others (*see*, Merrifield, *Biochemistry*, 3:1385-1390 (1964)), polyacrylamides, latex gels, polystyrene, dextran, polyacrylamides, rubber, silicon, plastics, nitrocellulose, celluloses, natural sponges. Of particular interest herein, are highly porous glasses (*see, e.g.*, U.S. Patent No. 4,244,721) and 20 others prepared by mixing a borosilicate, alcohol and water.

Synthetic supports include, but are not limited to: acrylamides, dextran-derivatives and dextran co-polymers, agarose-polyacrylamide blends, other polymers and co-polymers with various functional groups, methacrylate derivatives and co-polymers, polystyrene and polystyrene copolymers (*see, e.g.*, 25 Merrifield, *Biochemistry*, 3:1385-1390 (1964); Berg et al., in *Innovation Perspect. Solid Phase Synth. Collect. Pap.*, Int. Symp., 1st, Epton, Roger (Ed), pp. 453-459 (1990); Berg et al., *Pept., Proc. Eur. Pept. Symp.*, 20th, Jung, G. et al. (Eds), pp. 196-198 (1989); Berg et al., *J. Am. Chem. Soc.*, 111:8024-8026 (1989); Kent et al., *Isr. J. Chem.*, 17:243-247 (1979); Kent et al., *J. Org. Chem.*, 43:2845-2852 (1978); Mitchell et al., *Tetrahedron Lett.*, 42:3795-3798 (1976); U.S. Patent No. 4,507,230; U.S. Patent No. 4,006,117; 30 and U.S. Patent No. 5,389,449). Such materials include those made from

-117-

polymers and co-polymers such as polyvinylalcohols, acrylates and acrylic acids such as polyethylene-co-acrylic acid, polyethylene-co-methacrylic acid, polyethylene-co-ethylacrylate, polyethylene-co-methyl acrylate, polypropylene-co-acrylic acid, polypropylene-co-methyl-acrylic acid, polypropylene-co-ethylacrylate, 5 polypropylene-co-methyl acrylate, polyethylene-co-vinyl acetate, polypropylene-co-vinyl acetate, and those containing acid anhydride groups such as polyethylene-co-maleic anhydride and polypropylene-co-maleic anhydride. Liposomes have also been used as solid supports for affinity purifications (Powell et al. *Biotechnol. Bioeng.*, 33:173 (1989)).

- 10 Numerous methods have been developed for the immobilization of proteins and other biomolecules onto solid or liquid supports (*see, e.g.,* Mosbach, *Methods in Enzymology*, 44 (1976); Weetall, *Immobilized Enzymes, Antigens, Antibodies, and Peptides*, (1975); Kennedy et al., *Solid Phase Biochemistry, Analytical and Synthetic Aspects*, Scouten, ed., pp. 253-391 15 (1983); *see, generally*, Affinity Techniques. Enzyme Purification: *Part B. Methods in Enzymology*, Vol. 34, ed. W. B. Jakoby, M. Wilchek, Acad. Press, N.Y. (1974); and Immobilized Biochemicals and Affinity Chromatography, *Advances in Experimental Medicine and Biology*, vol. 42, ed. R. Dunlap, Plenum Press, N.Y. (1974)).
- 20 Among the most commonly used methods are absorption and adsorption or covalent binding to the support, either directly or via a linker, such as the numerous disulfide linkages, thioether bonds, hindered disulfide bonds, and covalent bonds between free reactive groups, such as amine and thiol groups, known to those of skill in art (*see, e.g.,* the PIERCE CATALOG, 25 ImmunoTechnology Catalog & Handbook, 1992-1993, which describes the preparation of and use of such reagents and provides a commercial source for such reagents; Wong, *Chemistry of Protein Conjugation and Cross Linking*, CRC Press (1993); *see also* DeWitt et al., *Proc. Natl. Acad. Sci. U.S.A.*, 90:6909 (1993); Zuckermann et al., *J. Am. Chem. Soc.*, 114:10646 (1992); Kurth et al., 30 *J. Am. Chem. Soc.*, 116:2661 (1994); Ellman et al., *Proc. Natl. Acad. Sci. U.S.A.*, 91:4708 (1994); Sucholeiki, *Tetrahedron Lttrs.*, 35:7307 (1994); Sun Wang, *J. Org. Chem.*, 41:3258 (1976); Padwa et al., *J. Org. Chem.*,

-118-

41:3550 (1971); and Vedejs et al., *J. Org. Chem.*, 49:575 (1984), which describe photosensitive linkers).

To effect immobilization, a composition containing the protein or other biomolecule is contacted with a support material such as alumina, carbon, an ion-exchange resin, cellulose, glass or a ceramic. Fluorocarbon polymers have been used as supports to which biomolecules have been attached by adsorption (see, U.S. Patent No. 3,843,443; Published International PCT Application WO/86 03840).

J. Prognosis and diagnosis

MTSP9 polypeptide proteins, domains, analogs, and derivatives thereof, and encoding nucleic acids (and sequences complementary thereto), and anti-MTSP9 polypeptide antibodies, can be used in diagnostics, particularly diagnosis of lung, head and neck, such as esophageal tumors, prostate, colon, ovary, cervix, breast and pancreas cancers. Such molecules can be used in assays, such as immunoassays, to detect, prognose, diagnose, or monitor various conditions, diseases, and disorders affecting MTSP9 polypeptide expression, or monitor the treatment thereof. For purposes herein, the presence of MTSP9s in body fluids or tumor tissues are of particular interest.

In particular, such an immunoassay is carried out by a method including contacting a sample derived from a patient with an anti-MTSP9 polypeptide antibody under conditions such that specific binding can occur, and detecting or measuring the amount of any specific binding by the antibody. Such binding of antibody, in tissue sections, can be used to detect aberrant MTSP9 polypeptide localization or aberrant (*e.g.*, increased, decreased or absent) levels of MTSP9 polypeptide. In a specific embodiment, antibody to an MTSP9 polypeptide can be used to assay in a patient tissue or body fluid, such as serum, sample for the presence of MTSP9 polypeptide where an aberrant level of MTSP9 polypeptide is an indication of a diseased condition.

The immunoassays which can be used include but are not limited to competitive and non-competitive assay systems using techniques such as western blots, radioimmunoassays, ELISA (enzyme linked immunosorbent assay), "sandwich" immunoassays, immunoprecipitation assays, precipitin

-119-

reactions, gel diffusion precipitin reactions, immunodiffusion assays, agglutination assays, complement-fixation assays, immunoradiometric assays, fluorescent immunoassays and protein A immunoassays.

MTSP9 polypeptide genes and related nucleic acid sequences and
5 subsequences, including complementary sequences, also can be used in hybridization assays. MTSP9 polypeptide nucleic acid sequences, or subsequences thereof containing about at least 8 nucleotides, generally 14 or 16 or 30 or more, generally less than 1000 or up to 100, contiguous nucleotides can be used as hybridization probes. Hybridization assays can be used to
10 detect, prognose, diagnose, or monitor conditions, disorders, or disease states associated with aberrant changes in MTSP9 polypeptide expression and/or activity as described herein. In particular, such a hybridization assay is carried out by a method by contacting a sample containing nucleic acid with a nucleic acid probe capable of hybridizing to MTSP9 polypeptide encoding DNA or RNA,
15 under conditions such that hybridization can occur, and detecting or measuring any resulting hybridization.

In a specific embodiment, a method of diagnosing a disease or disorder characterized by detecting an aberrant level of an MTSP9 polypeptide in a subject is provided herein by measuring the level of the DNA, RNA, protein or
20 functional activity of the MTSP9 polypeptide in a sample derived from the subject, wherein an increase or decrease in the level of the DNA, RNA, protein or functional activity of the MTSP9 polypeptide, relative to the level of the DNA, RNA, protein or functional activity found in an analogous sample not having the disease or disorder indicates the presence of the disease or disorder in the
25 subject.

Kits for diagnostic use are also provided, that contain in one or more containers an anti-MTSP9 polypeptide antibody, and, optionally, a labeled binding partner to the antibody. Alternatively, the anti-MTSP9 polypeptide antibody can be labeled (with a detectable marker, *e.g.*, a chemiluminescent,
30 enzymatic, fluorescent, or radioactive moiety). A kit is also provided that includes in one or more containers a nucleic acid probe capable of hybridizing to the MTSP9 polypeptide-encoding nucleic acid. In a specific embodiment, a kit

-120-

can include in one or more containers a pair of primers (*e.g.*, each in the size range of 6-30 nucleotides) that are capable of priming amplification [*e.g.*, by polymerase chain reaction (see *e.g.*, Innis et al., 1990, PCR Protocols, Academic Press, Inc., San Diego, CA), ligase chain reaction (see EP 320,308) use of Q β replicase, cyclic probe reaction, or other methods known in the art under appropriate reaction conditions of at least a portion of an MTSP9 polypeptide-encoding nucleic acid. A kit can optionally further include in a container a predetermined amount of a purified MTSP9 polypeptide or nucleic acid, *e.g.*, for use as a standard or control.

10 **K. Pharmaceutical compositions and modes of administration**

1. Components of the compositions

Pharmaceutical compositions containing the identified compounds that modulate the activity of an MTSP9 polypeptide are provided herein. Also provided are combinations of a compound that modulates the activity of an MTSP9 polypeptide and another treatment or compound for treatment of a neoplastic disorder, such as a chemotherapeutic compound.

The MTSP9 polypeptide modulator and the anti-tumor agent can be packaged as separate compositions for administration together or sequentially or intermittently. Alternatively, they can provided as a single composition for administration or as two compositions for administration as a single composition. The combinations can be packaged as kits.

a. MTSP9 polypeptide inhibitors

Any MTSP9 polypeptide inhibitors, including those described herein when used alone or in combination with other compounds, that can alleviate, reduce, ameliorate, prevent, or place or maintain in a state of remission of clinical symptoms or diagnostic markers associated with neoplastic diseases, including undesired and/or uncontrolled angiogenesis, can be used in the present combinations.

In one embodiment, the MTSP9 polypeptide inhibitor is an antibody or fragment thereof that specifically reacts with an MTSP9 polypeptide or the protease domain thereof, an inhibitor of the MTSP9 polypeptide production, an

-121-

inhibitor of MTSP9 polypeptide membrane-localization, or any inhibitor of the expression of or, especially, the activity of an MTSP9 polypeptide.

b. Anti-angiogenic agents and anti-tumor agents

Any anti-angiogenic agents and anti-tumor agents, including those
5 described herein, when used alone or in combination with other compounds, that can alleviate, reduce, ameliorate, prevent, or place or maintain in a state of remission of clinical symptoms or diagnostic markers associated with undesired and/or uncontrolled angiogenesis and/or tumor growth and metastasis, particularly solid neoplasms, vascular malformations and cardiovascular
10 disorders, crest syndromes, dermatological disorders, or ocular disorders, can be used in the combinations. Also contemplated are anti-tumor agents for use in combination with an inhibitor of an MTSP9 polypeptide.

c. Anti-tumor agents and anti-angiogenic agents

15 The compounds identified by the methods provided herein or provided herein can be used in combination with anti-tumor agents and/or anti-angiogenesis agents.

2. Formulations and route of administration

The compounds herein and agents can be formulated as pharmaceutical
20 compositions, typically for single dosage administration. The concentrations of the compounds in the formulations are effective for delivery of an amount, upon administration, that is effective for the intended treatment. Typically, the compositions are formulated for single dosage administration. To formulate a composition, the weight fraction of a compound or mixture thereof is dissolved,
25 suspended, dispersed or otherwise mixed in a selected vehicle at an effective concentration such that the treated condition is relieved or ameliorated. Pharmaceutical carriers or vehicles suitable for administration of the compounds provided herein include any such carriers known to those skilled in the art to be suitable for the particular mode of administration.

30 In addition, the compounds can be formulated as the sole pharmaceutically active ingredient in the composition or can be combined with other active ingredients. Liposomal suspensions, including tissue-targeted

-122-

liposomes, can also be suitable as pharmaceutically acceptable carriers. These can be prepared according to methods known to those skilled in the art. For example, liposome formulations can be prepared as described in U.S. Patent No. 4,522,811.

5 The active compound is included in the pharmaceutically acceptable carrier in an amount sufficient to exert a therapeutically useful effect in the absence of undesirable side effects on the patient treated. The therapeutically effective concentration can be determined empirically by testing the compounds in known in vitro and in vivo systems, such as the assays provided herein.

10 The concentration of active compound in the drug composition depends on absorption, inactivation and excretion rates of the active compound, the physicochemical characteristics of the compound, the dosage schedule, and amount administered as well as other factors known to those of skill in the art.

Typically a therapeutically effective dosage is contemplated. The
15 amounts administered can be on the order of 0.001 to 1 mg/ml, including about 0.005-0.05 mg/ml and about 0.01 mg/ml, of blood volume. Pharmaceutical dosage unit forms are prepared to provide from about 1 mg to about 1000 mg, including from about 10 to about 500 mg, and including about 25-75 mg of the essential active ingredient or a combination of essential ingredients per dosage
20 unit form. The precise dosage can be empirically determined.

The active ingredient can be administered at once, or can be divided into a number of smaller doses to be administered at intervals of time. It is understood that the precise dosage and duration of treatment is a function of the disease being treated and can be determined empirically using known testing
25 protocols or by extrapolation from in vivo or in vitro test data. It is to be noted that concentrations and dosage values can also vary with the severity of the condition to be alleviated. It is to be further understood that for any particular subject, specific dosage regimens should be adjusted over time according to the individual need and the professional judgment of the person administering or
30 supervising the administration of the compositions, and that the concentration ranges set forth herein are exemplary only and are not intended to limit the scope or use of the claimed compositions and combinations containing them.

Pharmaceutically acceptable derivatives include acids, salts, esters, hydrates, solvates and prodrug forms. The derivative is typically selected such that its pharmacokinetic properties are superior to the corresponding neutral compound.

5 Thus, effective concentrations or amounts of one or more of the compounds provided herein or pharmaceutically acceptable derivatives thereof are mixed with a suitable pharmaceutical carrier or vehicle for systemic, topical or local administration to form pharmaceutical compositions. Compounds are included in an amount effective for ameliorating or treating the disorder for
10 which treatment is contemplated. The concentration of active compound in the composition depends on absorption, inactivation, excretion rates of the active compound, the dosage schedule, amount administered, particular formulation as well as other factors known to those of skill in the art.

 Solutions or suspensions used for parenteral, intradermal, subcutaneous,
15 or topical application can include any of the following components: a sterile diluent, such as water for injection, saline solution, fixed oil, polyethylene glycol, glycerine, propylene glycol or other synthetic solvent; antimicrobial agents, such as benzyl alcohol and methyl parabens; antioxidants, such as ascorbic acid and sodium bisulfite; chelating agents, such as ethylenediaminetetraacetic acid
20 (EDTA); buffers, such as acetates, citrates and phosphates; and agents for the adjustment of tonicity such as sodium chloride or dextrose. Parenteral preparations can be enclosed in ampules, disposable syringes or single or multiple dose vials made of glass, plastic or other suitable material.

 In instances in which the compounds exhibit insufficient solubility,
25 methods for solubilizing compounds can be used. Such methods are known to those of skill in this art, and include, but are not limited to, using cosolvents, such as dimethylsulfoxide (DMSO), using surfactants, such as Tween®, or dissolution in aqueous sodium bicarbonate. Derivatives of the compounds, such as prodrugs of the compounds can also be used in formulating effective
30 pharmaceutical compositions. For ophthalmic indications, the compositions are formulated in an ophthalmically acceptable carrier. For the ophthalmic uses herein, local administration, either by topical administration or by injection are

-124-

contemplated. Time release formulations are also desirable. Typically, the compositions are formulated for single dosage administration, so that a single dose administers an effective amount.

Upon mixing or addition of the compound with the vehicle, the resulting mixture can be a solution, suspension, emulsion or other composition. The form of the resulting mixture depends upon a number of factors, including the intended mode of administration and the solubility of the compound in the selected carrier or vehicle. If necessary, pharmaceutically acceptable salts or other derivatives of the compounds are prepared.

The compound is included in the pharmaceutically acceptable carrier in an amount sufficient to exert a therapeutically useful effect in the absence of undesirable side effects on the patient treated. It is understood that number and degree of side effects depends upon the condition for which the compounds are administered. For example, certain toxic and undesirable side effects are tolerated when treating life-threatening illnesses that would not be tolerated when treating disorders of lesser consequence.

The compounds also can be mixed with other active materials, that do not impair the desired action, or with materials that supplement the desired action known to those of skill in the art. The formulations of the compounds and agents for use herein include those suitable for oral, rectal, topical, inhalational, buccal (*e.g.*, sublingual), parenteral (*e.g.*, subcutaneous, intramuscular, intradermal, or intravenous), transdermal administration or any route. The most suitable route in any given case depends on the nature and severity of the condition being treated and on the nature of the particular active compound which is being used. The formulations are provided for administration to humans and animals in unit dosage forms, such as tablets, capsules, pills, powders, granules, sterile parenteral solutions or suspensions, and oral solutions or suspensions, and oil-water emulsions containing suitable quantities of the compounds or pharmaceutically acceptable derivatives thereof. The pharmaceutically therapeutically active compounds and derivatives thereof are typically formulated and administered in unit-dosage forms or multiple-dosage forms. Unit-dose forms as used herein refers to physically discrete units suitable

-125-

for human and animal subjects and packaged individually as is known in the art. Each unit-dose contains a predetermined quantity of the therapeutically active compound sufficient to produce the desired therapeutic effect, in association with the required pharmaceutical carrier, vehicle or diluent. Examples of
5 unit-dose forms include ampoules and syringes and individually packaged tablets or capsules. Unit-dose forms can be administered in fractions or multiples thereof. A multiple-dose form is a plurality of identical unit-dosage forms packaged in a single container to be administered in segregated unit-dose form. Examples of multiple-dose forms include vials, bottles of tablets or capsules or
10 bottles of pints or gallons. Hence, multiple dose form is a multiple of unit-doses which are not segregated in packaging.

The composition can contain along with the active ingredient: a diluent such as lactose, sucrose, dicalcium phosphate, or carboxymethylcellulose; a lubricant, such as magnesium stearate, calcium stearate and talc; and a binder
15 such as starch, natural gums, such as gum acaciagelatin, glucose, molasses, polyvinylpyrrolidone, celluloses and derivatives thereof, povidone, crospovidones and other such binders known to those of skill in the art. Liquid pharmaceutically administrable compositions can, for example, be prepared by dissolving, dispersing, or otherwise mixing an active compound as defined above
20 and optional pharmaceutical adjuvants in a carrier, such as, for example, water, saline, aqueous dextrose, glycerol, glycols, ethanol, and the like, to thereby form a solution or suspension. If desired, the pharmaceutical composition to be administered can also contain minor amounts of nontoxic auxiliary substances such as wetting agents, emulsifying agents, or solubilizing agents, pH buffering
25 agents and the like, for example, acetate, sodium citrate, cyclodextrine derivatives, sorbitan monolaurate, triethanolamine sodium acetate, triethanolamine oleate, and other such agents. Methods of preparing such dosage forms are known, or will be apparent, to those skilled in this art (see, *e.g.*, Remington's Pharmaceutical Sciences, Mack Publishing Company, Easton,
30 Pa., 15th Edition, 1975). The composition or formulation to be administered contains a quantity of the active compound in an amount sufficient to alleviate the symptoms of the treated subject.

-126-

Dosage forms or compositions containing active ingredient in the range of 0.005% to 100% with the balance made up from non-toxic carrier can be prepared. For oral administration, the pharmaceutical compositions can take the form of, for example, tablets or capsules prepared by conventional means with

5 pharmaceutically acceptable excipients such as binding agents (*e.g.*, pregelatinized maize starch, polyvinyl pyrrolidone or hydroxypropyl methylcellulose); fillers (*e.g.*, lactose, microcrystalline cellulose or calcium hydrogen phosphate); lubricants (*e.g.*, magnesium stearate, talc or silica); disintegrants (*e.g.*, potato starch or sodium starch glycolate); or wetting agents

10 (*e.g.*, sodium lauryl sulphate). The tablets can be coated by methods well-known in the art.

The pharmaceutical preparation can also be in liquid form, for example, solutions, syrups or suspensions, or can be presented as a drug product for reconstitution with water or other suitable vehicle before use. Such liquid

15 preparations can be prepared by conventional means with pharmaceutically acceptable additives such as suspending agents (*e.g.*, sorbitol syrup, cellulose derivatives or hydrogenated edible fats); emulsifying agents (*e.g.*, lecithin or acacia); non-aqueous vehicles (*e.g.*, almond oil, oily esters, or fractionated vegetable oils); and preservatives (*e.g.*, methyl or propyl-p-hydroxybenzoates or

20 sorbic acid).

Formulations suitable for rectal administration can be presented as unit dose suppositories. These can be prepared by admixing the active compound with one or more conventional solid carriers, for example, cocoa butter, and then shaping the resulting mixture.

25 Formulations suitable for topical application to the skin or to the eye generally are formulated as an ointment, cream, lotion, paste, gel, spray, aerosol and oil. Carriers which can be used include vaseline, lanoline, polyethylene glycols, alcohols, and combinations of two or more thereof. The topical formulations can further advantageously contain 0.05 to 15 percent by weight

30 of thickeners selected from among hydroxypropyl methyl cellulose, methyl cellulose, polyvinylpyrrolidone, polyvinyl alcohol, poly (alkylene glycols), poly/hydroxyalkyl, (meth)acrylates or poly(meth)acrylamides. A topical

-127-

formulation is often applied by instillation or as an ointment into the conjunctival sac. It also can be used for irrigation or lubrication of the eye, facial sinuses, and external auditory meatus. It can also be injected into the anterior eye chamber and other places. The topical formulations in the liquid state can be
5 also present in a hydrophilic three-dimensional polymer matrix in the form of a strip, contact lens, and the like from which the active components are released.

For administration by inhalation, the compounds for use herein can be delivered in the form of an aerosol spray presentation from pressurized packs or a nebulizer, with the use of a suitable propellant, *e.g.*, dichlorodifluoromethane,
10 trichlorofluoromethane, dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol, the dosage unit can be determined by providing a valve to deliver a metered amount. Capsules and cartridges of, *e.g.*, gelatin, for use in an inhaler or insufflator can be formulated containing a powder mix of the compound and a suitable powder base such as
15 lactose or starch.

Formulations suitable for buccal (sublingual) administration include, for example, lozenges containing the active compound in a flavored base, usually sucrose and acacia or tragacanth; and pastilles containing the compound in an inert base such as gelatin and glycerin or sucrose and acacia.

20 The compounds can be formulated for parenteral administration by injection, *e.g.*, by bolus injection or continuous infusion. Formulations for injection can be presented in unit dosage form, *e.g.*, in ampules or in multi-dose containers, with an added preservative. The compositions can be suspensions, solutions or emulsions in oily or aqueous vehicles, and can contain formulatory
25 agents such as suspending, stabilizing and/or dispersing agents. Alternatively, the active ingredient can be in powder form for reconstitution with a suitable vehicle, *e.g.*, sterile pyrogen-free water or other solvents, before use.

Formulations suitable for transdermal administration can be presented as discrete patches adapted to remain in intimate contact with the epidermis of the
30 recipient for a prolonged period of time. Such patches suitably contain the active compound as an optionally buffered aqueous solution of, for example, 0.1 to 0.2 M concentration with respect to the active compound. Formulations

-128-

suitable for transdermal administration can also be delivered by iontophoresis (see, e.g., *Pharmaceutical Research* 3 (6), 318 (1986)) and typically take the form of an optionally buffered aqueous solution of the active compound.

The pharmaceutical compositions can also be administered by controlled
5 release means and/or delivery devices (see, e.g., in U.S. Patent Nos. 3,536,809; 3,598,123; 3,630,200; 3,845,770; 3,847,770; 3,916,899; 4,008,719; 4,687,610; 4,769,027; 5,059,595; 5,073,543; 5,120,548; 5,354,566; 5,591,767; 5,639,476; 5,674,533 and 5,733,566).

Desirable blood levels can be maintained by a continuous infusion of the
10 active agent as ascertained by plasma levels. It should be noted that the attending physician would know how to and when to terminate, interrupt or adjust therapy to lower dosage due to toxicity, or bone marrow, liver or kidney dysfunctions. Conversely, the attending physician would also know how to and when to adjust treatment to higher levels if the clinical response is not adequate
15 (precluding toxic side effects).

The efficacy and/or toxicity of the MTSP9 polypeptide inhibitor(s), alone or in combination with other agents also can be assessed by the methods known in the art (See generally, O'Reilly, *Investigational New Drugs*, 15:5-13 (1997)).

The active compounds or pharmaceutically acceptable derivatives can be
20 prepared with carriers that protect the compound against rapid elimination from the body, such as time release formulations or coatings.

Kits containing the compositions and/or the combinations with instructions for administration thereof are provided. The kit can further include a needle or syringe, typically packaged in sterile form, for injecting the complex,
25 and/or a packaged alcohol pad. Instructions are optionally included for administration of the active agent by a clinician or by the patient.

Finally, the compounds or MTSP9 polypeptides or protease domains thereof or compositions containing any of the preceding agents can be packaged as articles of manufacture containing packaging material, a compound or suitable
30 derivative thereof provided herein, which is effective for treatment of a diseases or disorders contemplated herein, within the packaging material, and a label that indicates that the compound or a suitable derivative thereof is for treating the

diseases or disorders contemplated herein. The label can optionally include the disorders for which the therapy is warranted.

L. Methods of treatment

The compounds identified by the methods herein are used for treating or preventing neoplastic diseases in an animal, particularly a mammal, including a human, is provided herein. In one embodiment, the method includes administering to a mammal an effective amount of an inhibitor of an MTSP9 polypeptide, whereby the disease or disorder is treated or prevented.

In an embodiment, the MTSP9 polypeptide inhibitor used in the treatment or prevention is administered with a pharmaceutically acceptable carrier or excipient. The mammal treated can be a human. The inhibitors provided herein are those identified by the screening assays. In addition, antibodies and antisense nucleic acids or double-stranded RNA (dsRNA), such as RNAi, are contemplated.

The treatment or prevention method can further include administering an anti-angiogenic treatment or agent or anti-tumor agent simultaneously with, prior to or subsequent to the MTSP9 polypeptide inhibitor, which can be any compound identified that inhibits the activity of an MTSP9 polypeptide. Such compounds include small molecule modulators, an antibody or a fragment or derivative thereof containing a binding region thereof against the MTSP9 polypeptide, an antisense nucleic acid or double-stranded RNA (dsRNA), such as RNAi, encoding an a portion of the MTSP9 polypeptide or complementary to thereto, and a nucleic acid containing at least a portion of a gene encoding the MTSP9 polypeptide into which a heterologous nucleotide sequence has been inserted such that the heterologous sequence inactivates the biological activity of at least a portion of the gene encoding the MTSP9 polypeptide, in which the portion of the gene encoding the MTSP9 polypeptide flanks the heterologous sequence to promote homologous recombination with a genomic gene encoding the MTSP9 polypeptide. In addition, such molecules are generally less than about 1000 nt long.

-130-

1. Antisense treatment

In a specific embodiment, as described hereinabove, MTSP9 polypeptide function is reduced or inhibited by MTSP9 polypeptide antisense nucleic acids, to treat or prevent neoplastic disease. The therapeutic or prophylactic use of nucleic acids of at least six nucleotides, generally up to about 150 nucleotides, that are antisense to a gene or cDNA encoding MTSP9 polypeptide or a portion thereof is provided. An MTSP9 polypeptide "antisense" nucleic acid as used herein refers to a nucleic acid capable of hybridizing to a portion of an MTSP9 polypeptide RNA (generally mRNA) by virtue of some sequence complementarity, and generally under high stringency conditions. The antisense nucleic acid can be complementary to a coding and/or noncoding region of an MTSP9 polypeptide mRNA. Such antisense nucleic acids have utility as therapeutics that reduce or inhibit MTSP9 polypeptide function, and can be used in the treatment or prevention of disorders as described *supra*.

The MTSP9 polypeptide antisense nucleic acids are of at least six nucleotides and are generally oligonucleotides (ranging from 6 to about 150 nucleotides including 6 to 50 nucleotides). The antisense molecule can be complementary to all or a portion of the protease domain. For example, the oligonucleotide is at least 10 nucleotides, at least 15 nucleotides, at least 100 nucleotides, or at least 125 nucleotides. The oligonucleotides can be DNA or RNA or chimeric mixtures or derivatives or modified versions thereof, single-stranded or double-stranded. The oligonucleotide can be modified at the base moiety, sugar moiety, or phosphate backbone. The oligonucleotide can include other appending groups such as peptides, or agents facilitating transport across the cell membrane (see, *e.g.*, Letsinger et al., *Proc. Natl. Acad. Sci. U.S.A.* 86:6553-6556 (1989); Lemaitre et al., *Proc. Natl. Acad. Sci. U.S.A.* 84:648-652 (1987); PCT Publication No. WO 88/09810, published December 15, 1988) or blood-brain barrier (see, *e.g.*, PCT Publication No. WO 89/10134, published April 25, 1988), hybridization-triggered cleavage agents (see, *e.g.*, Krol et al., *BioTechniques* 6:958-976 (1988)) or intercalating agents (see, *e.g.*, Zon, *Pharm. Res.* 5:539-549 (1988)).

-131-

The MTSP9 polypeptide antisense nucleic acid generally is an oligonucleotide, typically single-stranded DNA or RNA or an analog thereof or mixtures thereof. For example, the oligonucleotide includes a sequence antisense to a portion of a nucleic acid that encodes a human MTSP9

5 polypeptide. The oligonucleotide can be modified at any position on its structure with substituents generally known in the art.

The MTSP9 polypeptide antisense oligonucleotide can include at least one modified base moiety which is selected from the group including, but not limited to 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine,
 10 xanthine, 4-acetylcytosine, 5-(carboxyhydroxymethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine,
 15 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester,
 20 uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine.

In another embodiment, the oligonucleotide includes at least one modified sugar moiety selected from the group including but not limited to arabinose, 2-fluoroarabinose, xylulose, and hexose. The oligonucleotide can include at least
 25 one modified phosphate backbone selected from a phosphorothioate, a phosphorodithioate, a phosphoramidothioate, a phosphoramidate, a phosphordiamidate, a methylphosphonate, an alkyl phosphotriester, and a formacetal or analog thereof.

The oligonucleotide can be an α -anomeric oligonucleotide. An α -anomeric oligonucleotide forms specific double-stranded hybrids with complementary RNA
 30 in which the strands run parallel to each other (Gautier et al., *Nucl. Acids Res.* 15:6625-6641 (1987)).

-132-

The oligonucleotide can be conjugated to another molecule, such as, but are not limited to, a peptide, hybridization triggered cross-linking agent, transport agent or a hybridization-triggered cleavage agent. The oligonucleotides can be synthesized by standard methods known in the art, *e.g.* by use of an automated
5 DNA synthesizer (such as are commercially available from Biosearch, Applied Biosystems, etc.). As examples, phosphorothioate oligonucleotides can be synthesized by the method of Stein et al. (*Nucl. Acids Res.* 16:3209 (1988)), methylphosphonate oligonucleotides can be prepared by use of controlled pore glass polymer supports (Sarin et al., *Proc. Natl. Acad. Sci. U.S.A.* 85:7448-7451
10 (1988)), etc.

In a specific embodiment, the MTSP9 polypeptide antisense oligonucleotide includes catalytic RNA or a ribozyme (see, *e.g.*, PCT International Publication WO 90/11364, published October 4, 1990; Sarver et al., *Science* 247:1222-1225 (1990)). In another embodiment, the oligonucleotide is a 2'-O-methylribonucleotide (Inoue et al., *Nucl. Acids Res.* 15:6131-6148 (1987)), or a
15 chimeric RNA-DNA analogue (Inoue et al., *FEBS Lett.* 215:327-330 (1987)). Alternatively, the oligonucleotide can be double-stranded RNA (dsRNA) such as RNAi.

In an alternative embodiment, the MTSP9 polypeptide antisense nucleic
20 acid is produced intracellularly by transcription from an exogenous sequence. For example, a vector can be introduced *in vivo* such that it is taken up by a cell, within which cell the vector or a portion thereof is transcribed, producing an antisense nucleic acid (RNA). Such a vector would contain a sequence encoding the MTSP9 polypeptide antisense nucleic acid. Such a vector can remain
25 episomal or become chromosomally integrated, as long as it can be transcribed to produce the desired antisense RNA. Such vectors can be constructed by recombinant DNA technology methods standard in the art. Vectors can be plasmid, viral, or others known in the art, used for replication and expression in mammalian cells. Expression of the sequence encoding the MTSP9 polypeptide
30 antisense RNA can be by any promoter known in the art to act in mammalian, including human, cells. Such promoters can be inducible or constitutive. Such promoters include but are not limited to: the SV40 early promoter region

-133-

(Bernoist and Chambon, *Nature* 290:304-310 (1981), the promoter contained in the 3' long terminal repeat of Rous sarcoma virus (Yamamoto et al., *Cell* 22:787-797 (1980), the herpes thymidine kinase promoter (Wagner et al., *Proc. Natl. Acad. Sci. U.S.A.* 78:1441-1445 (1981), the regulatory sequences of the

5 metallothionein gene (Brinster et al., *Nature* 296:39-42 (1982), etc.

The antisense nucleic acids include sequence complementary to at least a portion of an RNA transcript of an MTSP9 polypeptide gene, including a human MTSP9 polypeptide gene. Absolute complementarity is not required.

The amount of MTSP9 polypeptide antisense nucleic acid that is effective

10 in the treatment or prevention of neoplastic disease depends on the nature of the disease, and can be determined empirically by standard clinical techniques. Where possible, it is desirable to determine the antisense cytotoxicity in cells *in vitro*, and then in useful animal model systems prior to testing and use in humans.

15 2. RNA interference

RNA interference (RNAi) (see, e.g. Chuang et al. (2000) *Proc. Natl. Acad. Sci. U.S.A.* 97:4985) can be employed to inhibit the expression of a gene encoding an MTSP9. Interfering RNA (RNAi) fragments, particularly double-stranded (ds) RNAi, can be used to generate loss-of-MTSP9 function. Methods

20 relating to the use of RNAi to silence genes in organisms including, mammals, C. elegans, Drosophila and plants, and humans are known (see, e.g., Fire et al. (1998) *Nature* 391:806-811; Fire (1999) *Trends Genet.* 15:358-363; Sharp (2001) *Genes Dev.* 15:485-490; Hammond, et al. (2001) *Nature Rev. Genet.* 2:110-1119; Tuschl (2001) *Chem. Biochem.* 2:239-245; Hamilton et al. (1999) *Science* 286:950-952; Hammond et al. (2000) *Nature* 404:293-296;

25 Zamore et al. (2000) *Cell* 101:25-33; Bernstein et al. (2001) *Nature* 409: 363-366; Elbashir et al. (2001) *Genes Dev.* 15:188-200; Elbashir et al. (2001) *Nature* 411:494-498; International PCT application No. WO 01/29058; International PCT application No. WO 99/32619).

30 Double-stranded RNA (dsRNA)-expressing constructs are introduced into a host, such as an animal or plant using, a replicable vector that remains episomal or integrates into the genome. By selecting appropriate sequences,

-134-

expression of dsRNA can interfere with accumulation of endogenous mRNA encoding an MTSP9. RNAi also can be used to inhibit expression *in vitro*. Regions include at least about 21 (or 21) nucleotides that are selective (i.e. unique) for MTSP9 are used to prepare the RNAi. Smaller fragments of about 21
5 nucleotides can be transformed directly (*i.e.*, *in vitro* or *in vivo*) into cells; larger RNAi dsRNA molecules are generally introduced using vectors that encode them. dsRNA molecules are at least about 21 bp long or longer, such as 50, 100, 150, 200 and longer. Methods, reagents and protocols for introducing nucleic acid molecules in to cells *in vitro* and *in vivo* are known to those of skill in the
10 art.

3. Gene Therapy

In an exemplary embodiment, nucleic acids that include a sequence of nucleotides encoding an MTSP9 polypeptide or functional domains or derivative thereof, are administered to promote MTSP9 polypeptide function, by way of
15 gene therapy. Gene therapy refers to therapy performed by the administration of a nucleic acid to a subject. In this embodiment, the nucleic acid produces its encoded protein that mediates a therapeutic effect by promoting MTSP9 polypeptide function. Any of the methods for gene therapy available in the art can be used (see, Goldspiel et al., *Clinical Pharmacy* 12:488-505 (1993); Wu
20 and Wu, *Biotherapy* 3:87-95 (1991); Tolstoshev, *An. Rev. Pharmacol. Toxicol.* 32:573-596 (1993); Mulligan, *Science* 260:926-932 (1993); and Morgan and Anderson, *An. Rev. Biochem.* 62:191-217 (1993); *TIBTECH* 11(5):155-215 (1993). For example, one therapeutic composition for gene therapy

includes an MTSP9 polypeptide-encoding nucleic acid that is part of an
25 expression vector that expresses an MTSP9 polypeptide or domain, fragment or chimeric protein thereof in a suitable host. In particular, such a nucleic acid has a promoter operably linked to the MTSP9 polypeptide coding region, the promoter being inducible or constitutive, and, optionally, tissue-specific. In another particular embodiment, a nucleic acid molecule is used in which the
30 MTSP9 polypeptide coding sequences and any other desired sequences are flanked by regions that promote homologous recombination at a desired site in the genome, thus providing for intrachromosomal expression of the SP protein

-135-

nucleic acid (Koller and Smithies, *Proc. Natl. Acad. Sci. USA* 86:8932-8935 (1989); Zijlstra et al., *Nature* 342:435-438 (1989)).

Delivery of the nucleic acid into a patient can be either direct, in which case the patient is directly exposed to the nucleic acid or nucleic acid-carrying
5 vector, or indirect, in which case, cells are first transformed with the nucleic acid *in vitro*, then transplanted into the patient. These two approaches are known, respectively, as *in vivo* or *ex vivo* gene therapy.

In a specific embodiment, the nucleic acid is directly administered *in vivo*, where it is expressed to produce the encoded product. This can be
10 accomplished by any of numerous methods known in the art, *e.g.*, by constructing it as part of an appropriate nucleic acid expression vector and administering it so that it becomes intracellular, *e.g.*, by infection using a defective or attenuated retroviral or other viral vector (see U.S. Patent No. 4,980,286), or by direct injection of naked DNA, or by use of microparticle
15 bombardment (*e.g.*, a gene gun; Biolistic, Dupont), or coating with lipids or cell-surface receptors or transfecting agents, encapsulation in liposomes, microparticles, or microcapsules, or by administering it in linkage to a peptide which is known to enter the nucleus, by administering it in linkage to a ligand subject to receptor-mediated endocytosis (see *e.g.*, Wu and Wu, *J. Biol. Chem.*
20 262:4429-4432 (1987)) (which can be used to target cell types specifically expressing the receptors), etc. In another embodiment, a nucleic acid-ligand complex can be formed in which the ligand is a fusogenic viral peptide to disrupt endosomes, allowing the nucleic acid to avoid lysosomal degradation. In yet another embodiment, the nucleic acid can be targeted *in vivo* for cell specific
25 uptake and expression, by targeting a specific receptor (see, *e.g.*, PCT Publications WO 92/06180 dated April 16, 1992 (Wu et al.); WO 92/22635 dated December 23, 1992 (Wilson et al.); WO92/20316 dated November 26, 1992 (Findeis et al.); WO93/14188 dated July 22, 1993 (Clarke et al.), WO 93/20221 dated October 14, 1993 (Young)). Alternatively, the nucleic acid can
30 be introduced intracellularly and incorporated within host cell DNA for expression, by homologous recombination (Koller and Smithies, *Proc. Natl. Acad. Sci. USA* 86:8932-8935 (1989); Zijlstra et al., *Nature* 342:435-438 (1989)).

-136-

In a specific embodiment, a viral vector that contains the MTSP9 polypeptide nucleic acid is used. For example, a retroviral vector can be used (see Miller et al., *Meth. Enzymol.* 217:581-599 (1993)). These retroviral vectors have been modified to delete retroviral sequences that are not necessary for packaging of the viral genome and integration into host cell DNA. The MTSP9 polypeptide nucleic acid to be used in gene therapy is cloned into the vector, which facilitates delivery of the gene into a patient. More detail about retroviral vectors can be found in Boesen et al., *Biotherapy* 6:291-302 (1994), which describes the use of a retroviral vector to deliver the *mdr1* gene to hematopoietic stem cells in order to make the stem cells more resistant to chemotherapy. Other references illustrating the use of retroviral vectors in gene therapy are: Clowes et al., *J. Clin. Invest.* 93:644-651 (1994); Kiem et al., *Blood* 83:1467-1473 (1994); Salmons and Gunzberg, *Human Gene Therapy* 4:129-141 (1993); and Grossman and Wilson, *Curr. Opin. in Genetics and Devel.* 3:110-114 (1993).

Adenoviruses are other viral vectors that can be used in gene therapy. Adenoviruses are especially attractive vehicles for delivering genes to respiratory epithelia. Adenoviruses naturally infect respiratory epithelia where they cause a mild disease. Other targets for adenovirus-based delivery systems are liver, the central nervous system, endothelial cells, and muscle. Adenoviruses have the advantage of being capable of infecting non-dividing cells. Kozarsky and Wilson, *Current Opinion in Genetics and Development* 3:499-503 (1993) present a review of adenovirus-based gene therapy. Bout et al., *Human Gene Therapy* 5:3-10 (1994) demonstrated the use of adenovirus vectors to transfer genes to the respiratory epithelia of rhesus monkeys. Other instances of the use of adenoviruses in gene therapy can be found in Rosenfeld et al., *Science* 252:431-434 (1991); Rosenfeld et al., *Cell* 68:143-155 (1992); and Mastrangeli et al., *J. Clin. Invest.* 91:225-234 (1993).

Adeno-associated virus (AAV) has also been proposed for use in gene therapy (Walsh et al., *Proc. Soc. Exp. Biol. Med.* 204:289-300 (1993)).

Another approach to gene therapy involves transferring a gene to cells in tissue culture by such methods as electroporation, lipofection, calcium

-137-

phosphate mediated transfection, or viral infection. Usually, the method of transfer includes the transfer of a selectable marker to the cells. The cells are then placed under selection to isolate those cells that have taken up and are expressing the transferred gene. Those cells are then delivered to a patient.

5 In this embodiment, the nucleic acid is introduced into a cell prior to administration *in vivo* of the resulting recombinant cell. Such introduction can be carried out by any method known in the art, including but not limited to transfection, electroporation, microinjection, infection with a viral or bacteriophage vector containing the nucleic acid sequences, cell fusion,
10 chromosome-mediated gene transfer, microcell-mediated gene transfer, spheroplast fusion, etc. Numerous techniques are known in the art for the introduction of foreign genes into cells (see *e.g.*, Loeffler and Behr, *Meth. Enzymol.* 217:599-618 (1993); Cohen et al., *Meth. Enzymol.* 217:618-644 (1993); Cline, *Pharmac. Ther.* 29:69-92 (1985)) and can be used, provided that
15 the necessary developmental and physiological functions of the recipient cells are not disrupted. The technique should provide for the stable transfer of the nucleic acid to the cell, so that the nucleic acid is expressible by the cell and generally heritable and expressible by its cell progeny.

 The resulting recombinant cells can be delivered to a patient by various
20 methods known in the art. In an embodiment, epithelial cells are injected, *e.g.*, subcutaneously. In another embodiment, recombinant skin cells can be applied as a skin graft onto the patient. Recombinant blood cells (*e.g.*, hematopoietic stem or progenitor cells) can be administered intravenously. The amount of cells envisioned for use depends on the desired effect, patient state, etc., and can be
25 determined by one skilled in the art.

 Cells into which a nucleic acid can be introduced for purposes of gene therapy encompass any desired, available cell type, and include but are not limited to epithelial cells, endothelial cells, keratinocytes, fibroblasts, muscle cells, hepatocytes; blood cells such as T lymphocytes, B lymphocytes,
30 monocytes, macrophages, neutrophils, eosinophils, megakaryocytes, granulocytes; various stem or progenitor cells, in particular hematopoietic stem

or progenitor cells, *e.g.*, such as stem cells obtained from bone marrow, umbilical cord blood, peripheral blood, fetal liver, and other sources thereof.

For example, a cell used for gene therapy is autologous to the patient. In an embodiment in which recombinant cells are used in gene therapy, an MTSP9
5 polypeptide nucleic acid is introduced into the cells such that it is expressible by the cells or their progeny, and the recombinant cells are then administered *in vivo* for therapeutic effect. In a specific embodiment, stem or progenitor cells are used. Any stem and/or progenitor cells which can be isolated and maintained *in vitro* can potentially be used in accordance with this embodiment.
10 Such stem cells include but are not limited to hematopoietic stem cells (HSC), stem cells of epithelial tissues such as the skin and the lining of the gut, embryonic heart muscle cells, liver stem cells (PCT Publication WO 94/08598, dated April 28, 1994), and neural stem cells (Stemple and Anderson, *Cell* 71:973-985 (1992)).

15 Epithelial stem cells (ESCs) or keratinocytes can be obtained from tissues such as the skin and the lining of the gut by known procedures (Rheinwald, *Meth. Cell Bio.* 21A:229 (1980)). In stratified epithelial tissue such as the skin, renewal occurs by mitosis of stem cells within the germinal layer, the layer closest to the basal lamina. Stem cells within the lining of the gut provide for a
20 rapid renewal rate of this tissue. ESCs or keratinocytes obtained from the skin or lining of the gut of a patient or donor can be grown in tissue culture (Rheinwald, *Meth. Cell Bio.* 21A:229 (1980); Pittelkow and Scott, *Cano Clinic Proc.* 61:771 (1986)). If the ESCs are provided by a donor, a method for suppression of host versus graft reactivity (*e.g.*, irradiation, drug or antibody
25 administration to promote moderate immunosuppression) also can be used.

With respect to hematopoietic stem cells (HSC), any technique which provides for the isolation, propagation, and maintenance *in vitro* of HSC can be used in this embodiment. Techniques by which this can be accomplished include (a) the isolation and establishment of HSC cultures from bone marrow
30 cells isolated from the future host, or a donor, or (b) the use of previously established long-term HSC cultures, which can be allogeneic or xenogeneic. Non-autologous HSC generally are used with a method of suppressing

-139-

transplantation immune reactions of the future host/patient. In a particular embodiment, human bone marrow cells can be obtained from the posterior iliac crest by needle aspiration (see, *e.g.*, Kodo et al., *J. Clin. Invest.* 73:1377-1384 (1984)). For example, the HSCs can be made highly enriched or in substantially
5 pure form. This enrichment can be accomplished before, during, or after long-term culturing, and can be done by any techniques known in the art. Long-term cultures of bone marrow cells can be established and maintained by using, for example, modified Dexter cell culture techniques (Dexter et al., *J. Cell Physiol.* 91:335 (1977) or Witlock-Witte culture techniques (Witlock and Witte, *Proc.*
10 *Natl. Acad. Sci. USA* 79:3608-3612 (1982)).

In a specific embodiment, the nucleic acid to be introduced for purposes of gene therapy includes an inducible promoter operably linked to the coding region, such that expression of the nucleic acid is controllable by controlling the presence or absence of the appropriate inducer of transcription.

15 3. Prodrugs

A method for treating tumors is provided. The method is practiced by administering a prodrug that is cleaved at a specific site by an MTSP9 to release an active drug or precursor that can be converted to active drug *in vivo*. Upon
20 contact with a cell that expresses MTSP9 activity, the prodrug is converted into an active drug. The prodrug can be a conjugate that contains the active agent, such as an anti-tumor drug, such as a cytotoxic agent, or other therapeutic agent (TA), linked to a substrate for the targeted MTSP9, such that the drug or agent is inactive or unable to enter a cell, in the conjugate, but is activated upon
25 cleavage. The prodrug, for example, can contain an oligopeptide, typically a relatively short, less than about 10 amino acids peptide, that is proteolytically cleaved by the targeted MTSP9. Cytotoxic agents, include, but are not limited to, alkylating agents, antiproliferative agents and tubulin binding agents. Others include, vinca drugs, mitomycins, bleomycins and taxanes.

-140-

M. Animal models

Transgenic animal models and animals, such as rodents, including mice and rats, cows, chickens, pigs, goats, sheep, monkeys, including gorillas, and other primates, are provided herein. In particular, transgenic non-human animals that contain heterologous nucleic acid encoding an MTSP9 polypeptide or a transgenic animal in which expression of the polypeptide has been altered, such as by replacing or modifying the promoter region or other regulatory region of the endogenous gene are provided. Such an animal can be produced by promoting recombination between endogenous nucleic acid and an exogenous MTSP9 gene that could be over-expressed or mis-expressed, such as by expression under a strong promoter, via homologous or other recombination event.

Transgenic animals can be produced by introducing the nucleic acid using any known method of delivery, including, but not limited to, microinjection, lipofection and other modes of gene delivery into a germline cell or somatic cells, such as an embryonic stem cell. Typically the nucleic acid is introduced into a cell, such as an embryonic stem cell (ES), followed by injecting the ES cells into a blastocyst, and implanting the blastocyst into a foster mother, which is followed by the birth of a transgenic animal. Generally introduction of a heterologous nucleic acid molecule into a chromosome of the animal occurs by a recombination between the heterologous MTSP9-encoding nucleic acid and endogenous nucleic acid. The heterologous nucleic acid can be targeted to a specific chromosome. In some instances, knockout animals can be produced. Such an animal can be initially produced by promoting homologous recombination between an MTSP9 polypeptide gene in its chromosome and an exogenous MTSP9 polypeptide gene that has been rendered biologically inactive (typically by insertion of a heterologous sequence, *e.g.*, an antibiotic resistance gene). In one embodiment, this homologous recombination is performed by transforming embryo-derived stem (ES) cells with a vector containing the insertionally inactivated MTSP9 polypeptide gene, such that homologous recombination occurs, followed by injecting the ES cells into a blastocyst, and implanting the blastocyst into a foster mother, followed by the birth of the

-141-

chimeric animal ("knockout animal") in which an MTSP9 polypeptide gene has been inactivated (see Capecchi, *Science* 244:1288-1292 (1989)). The chimeric animal can be bred to produce homozygous knockout animals, which can then be used to produce additional knockout animals. Knockout animals include, but
5 are not limited to, mice, hamsters, sheep, pigs, cattle, and other non-human mammals. For example, a knockout mouse is produced. The resulting animals can serve as models of specific diseases, such as cancers, that exhibit under-expression of an MTSP9 polypeptide. Such knockout animals can be used as animal models of such diseases *e.g.*, to screen for or test molecules for the
10 ability to treat or prevent such diseases or disorders.

Other types of transgenic animals also can be produced, including those that over-express the MTSP9 polypeptide. Such animals include "knock-in" animals that are animals in which the normal gene is replaced by a variant, such a mutant, an over-expressed form, or other form. For example, one species',
15 such as a rodent's endogenous gene can be replaced by the gene from an other species, such as from a human. Animals also can be produced by non-homologous recombination into other sites in a chromosome; including animals that have a plurality of integration events.

After production of the first generation transgenic animal, a chimeric
20 animal can be bred to produce additional animals with over-expressed or mis-expressed MTSP9 polypeptides. Such animals include, but are not limited to, mice, hamsters, sheep, pigs, cattle and other non-human mammals. The resulting animals can serve as models of specific diseases, such as cancers, that are exhibit over-expression or mis-expression of an MTSP9 polypeptide. Such
25 animals can be used as animal models of such diseases *e.g.*, to screen for or test molecules for the ability to treat or prevent such diseases or disorders. In a specific embodiment, a mouse with over-expressed or mis-expressed MTSP9 polypeptide is produced.

30 The following examples are included for illustrative purposes only and are not intended to limit the scope of the invention.

-142-

EXAMPLE 1**Identification of MTSP9**

The protein sequence of the protease domain of matriptase (MTSP1; accession number AF118224) was used to search the human HTGS (High
5 Throughput Genomic Sequence) database using the tblastn algorithm. This search and alignment algorithm compares a protein query sequence against a nucleotide sequence database dynamically translated in all six reading frames (both strands). Several potential serine proteases were identified, including one designated herein as MTSP9.

- 10 The translated sequence of MTSP9 has 36% identity to matriptase. MTSP9 appears to be localized in chromosome 15 (AC012571 clone). A search of sequences deposited in GenBank showed that no identical sequence has been deposited. Further search of the human EST database showed one EST clone (called HSU77054; GenBank accession number U77054) that matches almost
15 completely a short segment of MTSP9 sequence (nt 631 to nt 754 of SEQ ID No. 5, or the corresponding sequence in the full length clone nt. 1162-1279 of SEQ ID No. 17) except for 3 nucleotide mismatches.

Identification of tissue source for cloning of MTSP9

- Using the nucleotide sequence of MTSP9 derived from the genomic
20 sequence, two gene specific oligonucleotide primers were designed. The sequence for the 5' end primer is
5'-GGCAAGCTTCCCTTCAGTATGATAACATCCATCAG-3' (SEQ ID No. 7) and that of the 3' end primer is 5'-AATGAGATACCACGTATCTTTCAGATCCCTTG-3' (SEQ ID No. 8). These primers were used to screen a panel of 8 cDNA libraries
25 derived normal human tissues (Human Multiple Tissue cDNA Panel I; Clontech, Palo Alto, CA; catalog no. K1420-1). A band (~700 bp) was detected in human pancreas, and subsequent sequence analysis showed that the nucleotide sequence of this DNA fragment matched that of the genomic MTSP9 exon sequences.

30 **Gene expression profile of MTSP9 in normal, tumor tissues and cell lines**

To obtain a gene expression profile of the MTSP9 transcript, the MTSP9 cDNA fragment obtained from human pancreas was used to probe a dot blot

-143-

composed of RNA extracted from 76 different human tissues (Human Multiple Tissue Expression (MTE) Array; Clontech, Palo Alto, CA; catalog no. 7775-1). The results of this analysis indicate that MTSP9 is highly expressed in the esophagus and expressed at a low level in many other tissues. The MTSP9 transcript is found in kidney (adult and fetal), spleen (adult and fetal), placenta, liver (adult and fetal), thymus, peripheral blood leukocyte, lung (adult and fetal), pancreas, lymph node, bone marrow, trachea, uterus, prostate, esophagus, testes, ovary and the gland organs (mammary, adrenal, thyroid, pituitary and salivary). MTSP9 is also expressed in tumor esophagus tissues, in a lung carcinoma (A549 cell line) and, at a low level, in a colorectal carcinoma (SW480), lymphoma (Raji and Daudi), a cervical carcinoma (HeLaS3) and leukemia (HL-60, K-562 and MOLT-4) cell lines.

PCR amplification of cDNA encoding full-length protease domain of MTSP9

To obtain the cDNA fragment encoding the protease domain of MTSP9, an end-to-end PCR amplification using gene-specific primers and the cDNA library from human esophagus was used. The two primers used were: 5'-CGAGTTGTTCCATTAAACGTCAACAGAATAGC-3' (SEQ ID No. 9) for the 5' end and 5'-GCATACAGCTTTCTTTGTTTAACTTTTATCGTG -3' (SEQ ID No. 10) for the 3' end. The sequences for both primers were derived from the genomic sequence of MTSP9. The 5' primer contained the sequence that encodes a region immediately upstream of the start of the MTSP9 protease domain (RVVPLNVNRIA; SEQ ID NO. 12). The 3' primer corresponds to the sequence immediately after the presumed stop codon. A ~750-bp fragment was amplified from the human esophagus cDNA library. The PCR product was isolated and purified using the QIAquick gel extraction kit (Qiagen, Valencia, CA; catalog no. 28704). The MTSP9 PCR product was used to amplify the cDNA fragment containing the appropriate restriction sites for subcloning into the Pichia vector, pPIC9K. The gene-specific primers used were 5'-TCTCTCGAGAAAAGAATAGCATCTGGAGTCATTGCACCCAAG-3' (SEQ ID No. 13) at the 5' end and 5'-ATAGCGGCCGCATTAGATGCCTGTTTTTGAAGCAATC-3' (SEQ ID No. 14) at the 3' end. The 5' end primer contained an XhoI site (underlined) immediately

-144-

upstream of the Pichia protease cleavage site and part of the MTSP9 protease domain (KRIASGVIAPK; SEQ ID No. 15), while the 3' end primer contained a NotI site (underlined) immediately downstream of the stop codon (in bold).

Cloning of full-length cDNA of MTSP9 by RACE

- 5 RACE-ready cDNA libraries were prepared from both normal and tumor human esophagus poly A + RNAs using SMART-RACE cDNA amplification kit (Clontech; catalog no. K1811-1) and First Choice RLM-RACE kit (Ambion, Austin, TX; catalog no. 1700). For 5'-RACE reaction, an anti-sense gene-specific primer
- 10 (5'- AATGAGATACACGTATCTTTCAGATCCCTTG-3' SEQ ID No. 19) together with the sense primer hybridizing to the adapter present at the 5' ends of the cDNAs was used. A band (~1.3 kbp) was amplified and identified by Southern analysis against a probe consisting of the protease domain of MTSP9. The 3'-RACE reaction was done in a similar fashion except that the sense
- 15 gene-specific primer used was 5'-ATGAGAAGTACCGCTCTGCAGCAAGAGAG-3' (SEQ ID No. 20). A band (~0.8 kbp) was amplified and isolated from agarose gel. The two RACE products were separately subcloned into E. coli vectors using TA cloning (TOPO TA cloning kit; Invitrogen, Carlsbad, CA; catalog no. K4500-01). After
- 20 transformation, plasmid DNAs from representative clones were isolated, purified and digested with EcoRI to check for the presence of inserts. The plasmid DNAs were sequenced initially with M13 forward and reverse primers, then by gene-specific primers spanning the whole insert in both directions.

Serine protease domain of MTSP9 and homology to other proteases

- 25 Sequence analysis of the translated coding region of MTSP9 showed the presence of a transmembrane domain at the N terminus and a trypsin-like serine protease domain at the C terminus. In between these domains is a stretch of protein sequence (149 amino acid residues long) with no known discernible domain, and which shares 20% identity with the same stretch of protein
- 30 sequence in endothelase 1. Overall, the full-length protein sequence of MTSP9 shares 42% identity with human endothelase 1 (DESC1; Genbank accession number AF064819) and 40% identity with another type II membrane type serine

-145-

protease, human airway trypsin-like serine protease (Genbank accession number NP004253). Analysis of the MTSP9 protease domain sequence indicated that it is a trypsin-like serine protease domain characterized by the presence of a protease activation cleavage site at the beginning of the domain and the catalytic triad residues (histidine, aspartate and serine) in 3 highly-conserved regions of the catalytic domain. Alignment of the protease domain sequence showed 56% identity with that of endotheliase 1 and 48% identity with the human airway trypsin-like protease domain.

Sequence analysis

MTSP9 cDNA and protein sequences were analyzed using MacVector (version 6.5; Oxford Molecular Ltd., Madison, WI). The cDNA encoding the protease domain of MTSP9 is 699 bp long which translates to a 232-amino acid protein. The nucleotide sequence of the protease domain and the translated protein sequence of MTSP9 are as follows (see SEQ ID Nos. 5, 6 and 16):

MTSP9-cDNA containing protease domain sequence Range: 1 to 777 (protease encoding domain 31-729)

```

      10      20      30      40      50      60
20 AAACGAGTTGTTCCATTAAACGTCAACAGAAATAGCATCTGGAGTCATTGCACCCAAGGCG
   TTTGCTCAACAAGGTAATTTGCAGTTGTCTTATCGTAGACCTCAGTAACGTGGGTTCGCG
      70      80      90     100     110     120
   GCCTGGCCTTGGCAAGCTTCCCTTCAGTATGATAACATCCATCAGTGTGGGGCCACCTTG
   CGGACCGGAACCGTTTCAAGGGAAGTCATACTATTGTAGGTAGTCACACCCCGGTGGAAC
      130     140     150     160     170     180
25 ATTAGTAACACATGGCTTGTCACTGCAGCACACTGCTTCCAGAAGTATAAAAATCCACAT
   TAATCATTGTGTACCGAACAGTGACGTCGTGTGACGAAGGTCTTCATATTTTATAGGTGTA
      190     200     210     220     230     240
   CAATGGACTGTTAGTTTTTGGAAACAAAATCAACCTCCCTTAATGAAAAGAAATGTCAGA
   GTTACCTGACAATCAAAACCTTGTTTTTAGTTGGGAGGGAATTACTTTTCTTTACAGTCT
      250     260     270     280     290     300
30 AGATTTATTATCCATGAGAAGTACCGCTCTGCAGCAAGAGAGTACGACATTGCTGTTGTG
   TCTAAATAATAGGTACTCTTCATGGCGAGACGTCGTTCTCTCATGCTGTAACGACAACAC
      310     320     330     340     350     360
   CAGGTCTCTTCCAGAGTCACCTTTTTCGGATGACATACGCCGGATTGTTTGCCAGAAGCC
35 GTCCAGAGAAGGTCTCAGTGGAAAAGCCTACTGTATGCGGCCTAAACAAACGGTCTTCGG
      370     380     390     400     410     420
   TCTGCATCCTTCCAACCAAATTTGACTGTCCACATCACAGGATTTGGAGCACTTTACTAT
   AGACGTAGGAAGGTTGGTTTAAACTGACAGGTGTAGTGTCTTAAACCTCGTGAAATGATA
      430     440     450     460     470     480
40 GGTGGGGAATCCCAAAATGATCTCCGAGAAGCCAGAGTGAAAATCATAAGTGACGATGTC
   CCACCCCTTAGGGTTTTACTAGAGGCTCTTCGGTCTCACTTTTAGTATTCACTGCTACAG
      490     500     510     520     530     540
   TGCAAGCAACCACAGGTGTATGGCAATGATATAAAACCTGGAATGTTCTGTGCCGGATAT
   ACGTTCGTTGGTGTCCACATACCGTTACTATATTTTGGACCTTACAAGACACGGCCTATA
      550     560     570     580     590     600
45 ATGGAAGGAATTTATGATGCCTGCAGGGGTGATTCTGGGGGACCTTTAGTCACAAGGGAT
   TACCTTCCTTAAATACTACGGACGTCCCACTAAGACCCCTGGAATCAGTGTTCCTTA
      610     620     630     640     650     660
   CTGAAAGATACGTGGTATCTCATTGGAATTGTAAGCTGGGGAGATAACTGTGGTCAAAAG

```

-146-

GACTTTCTATGCACCATAGAGTAACCTTAACATTTCGACCCCTCTATTGACACCAGTTTTTC
 670 680 690 700 710 720
 GACAAGCCTGGAGTCTACACACAAGTGACTTATTACCGAAACTGGATTGCTTCAAAAACA
 CTGTTCCGACCTCAGATGTGTGTTCACTGAATAATGGCTTTGACCTAACGAAGTTTTTGT
 5 730 740 750 760 770
 GGCATCTAATTCACGATAAAAGTTAAACAAAGAAAGCTGTATGCAGGTCATATATGC
 CCGTAGATTAAGTGCTATTTTCAATTTGTTTCTTTTCGACATACGTCCAGTATATACG

10 MTSP9-protease domain cDNA & protein sequence Range: 31 to 729

10 20 30 40 50 60
 AAACGAGTTGTTCCATTAAACGTCAACAGAATAGCATCTGGAGTCATTGCACCCAAGGCG
 TTTGCTCAACAAGGTAATTTGCAGTTGTCTTATCGTAGACCTCAGTAACGTGGGTTCGCG
 15 K R V V P L N V N R I A S G V I A P K A
 70 80 90 100 110 120
 GCCTGGCCTTGGCAAGCTTCCCTTCAGTATGATAACATCCATCAGTGTGGGGCCACCTTG
 CGGACCGGAACCGTTTCAAGGAAGTCATACCTATTGTAGGTAGTCACACCCCGGTGGAAC
 20 A W P W Q A S L Q Y D N I H Q C G A T L
 130 140 150 160 170 180
 ATTAGTAACACATGGCTTGTCACTGCAGCACACTGCTTCCAGAGTATAAAAATCCACAT
 TAATCATTGTGTACCGAACAGTGACGTCGTGTGACGAAGGTCTTCATATTTTTAGGTGTA
 25 I S N T W L V T A A H C F Q K Y K N P H
 190 200 210 220 230 240
 CAATGGACTGTTAGTTTTTGAACAAAAATCAACCCTCCCTTAATGAAAAGAAATGTCAGA
 GTTACCTGACAATCAAAACCTTGTTTTTAGTTGGGAGGGAATTACTTTTCTTACAGTCT
 30 Q W T V S F G T K I N P P L M K R N V R
 250 260 270 280 290 300
 AGATTTATTATCCATGAGAAGTACCGCTCTGCAGCAAGAGAGTACGACATTGCTGTTGTG
 TCTAAATAATAGGTACTCTTCATGGCGAGACGTCGTTCTCTCATGCTGTAACGACAACAC
 35 R F I I H E K Y R S A A R E Y D I A V V
 310 320 330 340 350 360
 CAGGTCTCTTCCAGAGTCACCTTTTTCGGATGACATACGCCGATTGTTTGCCAGAAGCC
 GTCCAGAGAAGGTCTCAGTGGAAAAGCCTACTGTATGCGGCCTAAACAAACGGTCTTTCGG
 40 Q V S S R V T F S D D I R R I C L P E A
 370 380 390 400 410 420
 TCTGCATCCTTCCAACCAAATTTGACTGTCCACATCACAGGATTTGGAGCACTTTACTAT
 AGACGTAGGAAGGTTGGTTTAAACTGACAGGTGTAGTGTCTTAAACCTCGTGAAATGATA
 45 S A S F Q P N L T V H I T G F G A L Y Y
 430 440 450 460 470 480
 GGTGGGAATCCCAAAATGATCTCCGAGAAGCCAGAGTGAAAATCATAAGTGACGATGTC
 CCACCCCTTAGGGTTTTACTAGAGGCTCTTCGGTCTCACTTTTAGTATTCACTGCTACAG
 50 G G E S Q N D L R E A R V K I I S D D V
 490 500 510 520 530 540
 TGCAAGCAACCACAGGTGTATGGCAATGATATAAAACCTGGAATGTTCTGTGCCGGATAT
 ACGTTTCGTTGGTGTCCACATACCGTTACTATATTTTGGACCTTACAAGACACGGCCTATA
 55 C K Q P Q V Y G N D I K P G M F C A G Y
 550 560 570 580 590 600
 ATGGAAGGAATTTATGATGCCTGCAGGGGTGATTCTGGGGGACCTTTAGTCACAAGGGAT
 TACCTTCCTTAAATACTACGGACGTCCCCACTAAGACCCCTGGAATCAGTGTTCCTTA
 60 M E G I Y D A C R G D S G G P L V T R D
 610 620 630 640 650 660

-147-

```

CTGAAAGATACGTGGTATCTCATTGGAATTGTAAGCTGGGGAGATAACTGTGGTCAAAAG
GACTTTCTATGCACCATAGAGTAACCTTAACATTCGACCCCTCTATTGACACCAGTTTTTC
  L  K  D  T  W  Y  L  I  G  I  V  S  W  G  D  N  C  G  Q  K
5
      670      680      690      700      710      720
GACAAGCCTGGAGTCTACACACAAGTGACTTATTACCGAAACTGGATTGCTTCAAAAACA
CTGTTTCGGACCTCAGATGTGTGTTCCTGAATAATGGCTTTGACCTAACGAAGTTTTTGT
  D  K  P  G  V  Y  T  Q  V  T  Y  Y  R  N  W  I  A  S  K  T
10
GGCATCTAA
CCGTAGATT
  G  I  *

```

MTSP9 cDNA and protein sequences were analyzed using MacVector
 15 (version 6.5; Oxford Molecular Ltd., Madison, WI). The full-length encoding
 clone is 1,422 bp long, with a coding region of 1,257 bp long. The translated
 protein sequence is 418 amino acid residues long. The DNA encoding the
 protease domain of MTSP9 is 699 bp long which translate to a 232-amino acid
 protein.

20 MTSP9 full-length cDNA sequence
 Sequence Range: 1 to 1422

```

      10      20      30      40      50      60
25 GCGAGCTGGAAGGACGAATCTCCGGCAGCATTTCATTACGACAAATGAATGCTGCCGGAT
   CCGCTCGACCTTCCTGCTTAGAGGCCGTCGTAAGTAATGCTGTTTACTTACGACGGCCTA

      70      80      90      100     110     120
   TAGAAAGTTGAGTTTCAGTGGGTGCAGACCTGCAAGATCATATTCTTCCTCCTGTACATGA
   ATCTTTCAACTCAAGTCACCCACGTCTGGACGTTCTAGTATAAGAAGGAGGACATGTACT

30      130     140     150     160     170     180
   TGTATCGGACAGTAGGATTTGGCACCCGAAGCAGAAATCTGAAGCCATGGATGATTGCCG
   ACATAGCCTGTTCATCCTAAACCGTGGGCTTCGTCCTTTAGACTTCGGTACCTACTAACGGC

35      190     200     210     220     230     240
   TTCTCATTGTGTTGTCCCTGACAGTGGTGGCAGTGACCATAGGTCTCCTGGTTCACTTCC
   AAGAGTAACACAACAGGGACTGTCAACACCGTCACTGGTATCCAGAGGACCAAGTGAAGG

40      250     260     270     280     290     300
   TAGTATTTGACCAAAAAAGGAGTACTATCATGGCTCCTTTAAAATTTTAGATCCACAAA
   ATCATAAACTGGTTTTTTTCCTCATGATAGTACCGAGGAAATTTTAAAATCTAGGTGTTT

45      310     320     330     340     350     360
   TCAATAACAATTTTCGGACAAAGCAACACATATCAACTTAAGGACTTACGAGAGACGACCG
   AGTTATTGTTAAAGCCTGTTTCGTTGTGTATAGTTGAATTCCTGAATGCTCTCTGCTGGC

50      370     380     390     400     410     420
   AAAATTTGGTGGATGAGATATTTATAGATTTCAGCCTGGAAGAAAAATTATATCAAGAACC
   TTTTAAACCACCTACTCTATAAATATCTAAGTCGGACCTTCTTTTAAATATAGTTCTTGG

55      430     440     450     460     470     480
   AAGTAGTCGACTGACTCCAGAGGAAGATGGTGTGAAAGTAGATGTCATTATGGTGTTC
   TTCATCAGTCTGACTGAGGTCTCCTTCTACCACACTTTCATCTACAGTAATACCACAAGG

      490     500     510     520     530     540
   AGTTCCCCTCTACTGAACAAAGGGCAGTAAGAGAGAAGAAAATCCAAAGCATCTTAAATC

```


TCAAGGGGAGATGACTTGTTCCCGTCATTCTCTCTTCTTTTAGGTTTCGTAGAATTTAG

[illegible]

-149-

MTSP9 full-length coding sequence

Sequence Range: 1 to 1257

```

5      10      20      30      40      50      60
   ATGATGTATCGGACAGTAGGATTTGGCACCCGAAGCAGAAATCTGAAGCCATGGATGATT
   TACTACATAGCCTGTCATCCTAAACCGTGGGCTTCGTCTTTAGACTTCGGTACCTACTAA

10      70      80      90     100     110     120
   GCCGTTCTCATTGTGTTGTCCCTGACAGTGGTGGCAGTGACCATAGGTCTCCTGGTTTAC
   CGGCAAGAGTAACACAACAGGGACTGTCACCACCGTCAC TGGTATCCAGAGGACCAAGTG

15     130     140     150     160     170     180
   TTCCTAGTATTTGACCAAAAAAGGAGTACTATCATGGCTCCTTTAAAATTTTAGATCCA
   AAGGATCATAAACTGGTTTTTTTCCCTCATGATAGTACCGAGGAAATTTTAAAATCTAGGT

20     190     200     210     220     230     240
   CAAATCAATAACAATTTTCGGACAAAGCAACACATATCAACTTAAGGACTTACGAGAGACG
   GTTTAGTTATTGTTAAAGCCTGTTTCGTTGTGTATAGTTGAATTCCTGAATGCTCTCTGC

25     250     260     270     280     290     300
   ACCGAAAATTTGGTGGATGAGATATTTATAGATTCAGCCTGGAAGAAAAATTATATCAAG
   TGGCTTTTAAACCACCTACTCTATAAATATCTAAGTCGGACCTTCTTTTAAATATAGTTC

30     310     320     330     340     350     360
   AACCAAGTAGTCAGACTGACTCCAGAGGAAGATGGTGTGAAAGTAGATGTCATTATGGTG
   TTGGTTCATCAGTCTGACTGAGGTCTCCTTCTACCACACTTTCATCTACAGTAATACCAC

35     370     380     390     400     410     420
   TTCCAGTTCCCCTCTACTGAACAAAGGGCAGTAAGAGAGAAGAAAATCCAAAGCATCTTA
   AAGGTCAAGGGGAGATGACTTGTTTCCCGTCATTCTCTCTTCTTTTAGGTTTCGTAGAAT

40     430     440     450     460     470     480
   AATCAGAAGATAAGGAATTTAAGAGCCTTGCCAATAAATGCCTCATCAGTTCAAGTTAAT
   TTAGTCTTCTATTCTTAAATTCTCGGAACGGTTATTTACGGAGTAGTCAAGTTCAATTA

45     490     500     510     520     530     540
   GCAATGAGCTCATCAACAGGGGAGTTAACTGTCCAAGCAAGTTGTGGTAAACGAGTTGTT
   CGTTACTCGAGTAGTTGTCCCTCAATTGACAGGTTTCGTTCAACACCATTGTGCTCAACAA

50     550     560     570     580     590     600
   CCATTAAACGTCAACAGAATAGCATCTGGAGTCATTGCACCCAAGGCGGCCTGGCCTTGG
   GGTAAATTTGCAGTTGTCTTATCGTAGACCTCAGTAACGTGGGTTCCGCCGGACCGGAACC

55     610     620     630     640     650     660
   CAAGCTTCCCTTCAGTATGATAACATCCATCAGTGTGGGGCCACCTTGATTAGTAACACA
   GTTCGAAGGGAAGTCATACTATTGTAGGTAGTCACACCCCGGTGGAATAATCATTGTGT

60     670     680     690     700     710     720
   TGGCTTGTCAGTGCAGCACACTGCTTCCAGAAGTATAAAAAATCCACATCAATGGACTGTT
   ACCGAACAGTGACGTCGTGTGACGAAGGTCTTCATATTTTTAGGTGTAGTTACCTGACAA

65     730     740     750     760     770     780
   AGTTTTGGAACAAAAATCAACCCTCCCTTAATGAAAAGAAATGTCAGAAGATTTATTATC
   TCAAAACCTTGTTTTTAGTTGGGAGGGAATTACTTTTCTTTACAGTCTTCTAAATAATAG

70     790     800     810     820     830     840
   CATGAGAAGTACCGCTCTGCAGCAAGAGAGTACGACATTGCTGTTGTGTCAGGTCTCTTCC
   GTACTCTTCATGGCGAGACGTCGTTCTCTCATGCTGTAACGACAACACGTCCAGAGAAGG

75     850     860     870     880     890     900
   AGAGTCACCTTTTCGGATGACATACGCCGATTGTTTGCCAGAAGCCTCTGCATCCTTC
   TCTCAGTGGAAAAGCCTACTGTATGCGGCCATAACAAACGGTCTTCGGAGACGTAGGAAG
```

-150-

910 920 930 940 950 960
 CAACCAAATTTGACTGTCCACATCACAGGATTTGGAGCACTTTACTATGGTGGGGAATCC
 GTTGGTTTAACTGACAGGTGTAGTGTCTTAAACCTCGTGAATGATACCACCCCTTAGG
 5 970 980 990 1000 1010 1020
 CAAATGATCTCCGAGAAGCCAGAGTGAAATCATAAGTGACGATGTCTGCAAGCAACCA
 GTTTTACTAGAGGCTCTTCGGTCTCACTTTTAGTATTCAGTGTACAGACGTTTCGTTGGT
 10 1030 1040 1050 1060 1070 1080
 CAGGTGTATGGCAATGATATAAAACCTGGAATGTTCTGTGCCGGATATATGGAAGGAATT
 GTCCACATACCGTTACTATATTTTGGACCTTACAAGACACGGCCTATATACCTTCCTTAA
 15 1090 1100 1110 1120 1130 1140
 TATGATGCCTGCAGGGGTGATTCTGGGGGACCTTTAGTCACAAGGGATCTGAAAGATACG
 ATACTACGGACGTCCCACTAAGACCCCTTGAAATCAGTGTTCCTTAGACTTTCTATGC
 20 1150 1160 1170 1180 1190 1200
 TGGTATCTCATTGGAATTGTAAGCTGGGGAGATAACTGTGGTCAAAAGGACAAGCCTGGA
 ACCATAGAGTAACCTTAACATTCGACCCCTCTATTGACACCAGTTTTCTGTTCGGACCT
 25 1210 1220 1230 1240 1250
 GTCTACACACAAGTGACTTATTACCGAAACTGGATTGCTTCAAAAACAGGCATCTAA
 CAGATGTGTGTTCACTGAATAATGGCTTTGACCTAACGAAGTTTTGTCCGTAGATT

MTSP9 full-length protein sequence

Sequence Range: 1 to 419

10 20 30 40 50 60
 MMYRTVGFGRSRNLKPWMIAVLIVLSLTVVAVTIGLLVHFLVFDQKKEYYHGSFKILDP
 _____ 1 TO 1257 OF MTSP9/CODING/CDNA/FINAL (TRANSLATED) _____>
 70 80 90 100 110 120
 QINNNFGQSNTRYQLKDLRETTENLVDEIFIDSAWKKNYIKNQVRLTPEEDGVKVDVIMV
 _____ 1 TO 1257 OF MTSP9/CODING/CDNA/FINAL (TRANSLATED) _____>
 35 130 140 150 160 170 180
 FQFPSTEQRAVREKKIQSILNQKIRNLRALPINASSVQVNAMSSSTGELTVQASCGKRVV
 _____ 1 TO 1257 OF MTSP9/CODING/CDNA/FINAL (TRANSLATED) _____>
 40 190 200 210 220 230 240
 PLNVNRIASGVIAPKAAWPQASLQYDNIHQCGATLISNTWLVTAAHCFQKYKNPHQWTV
 _____ 1 TO 1257 OF MTSP9/CODING/CDNA/FINAL (TRANSLATED) _____>
 45 250 260 270 280 290 300
 SFGTKINPPLMKRNVRRFIIHEKYRSAAREYDIAVVQVSSRVTFSDDIRRICLPEASASF
 _____ 1 TO 1257 OF MTSP9/CODING/CDNA/FINAL (TRANSLATED) _____>
 50 310 320 330 340 350 360
 QPNLTVHITGFGALYYGGESQNDLREARVKIISDDVCKQPQVYGNDIKPGMFCAGYMEGI
 _____ 1 TO 1257 OF MTSP9/CODING/CDNA/FINAL (TRANSLATED) _____>
 55 370 380 390 400 410
 YDACRGDSGGPLVTRDLKDTWYLIGTVSWGDNCGQKDKPGVYTQVTYYRNWIASKTGI*
 _____ 1 TO 1257 OF MTSP9/CODING/CDNA/FINAL (TRANSLATED) _____>

-151-

EXAMPLE 2

Expression of the protease MTSP domains

Nucleic acid encoding each the MTSP9 and protease domain thereof was cloned into a derivative of the *Pichia pastoris* vector pPIC9K (available from
5 Invitrogen; see SEQ ID NO. 11). Plasmid pPIC9K features include the 5' AOX1 promoter fragment at 1-948; 5' AOX1 primer site at 855-875; alpha-factor secretion signal(s) at 949-1218; alpha-factor primer site at 1152-1172; multiple cloning site at 1192-1241; 3' AOX1 primer site at 1327-1347; 3' AOX1 transcription termination region at 1253-1586; HIS4 ORF at 4514-1980;
10 kanamycin resistance gene at 5743-4928; 3' AOX1 fragment at 6122-6879; ColE1 origin at 7961-7288; and the ampicillin resistance gene at 8966-8106. The plasmid is derived from pPIC9K by eliminating the XhoI site in the kanamycin resistance gene and the resulting vector is herein designated pPIC9Kx.

15 C122S mutagenesis of the Protease domain of MTSP9

The gene encoding the protease domain of MTSP9 was mutagenized by PCR SOE (PCR-based splicing by overlap extension) to replace the unpaired cysteine at position 122 (chymotrypsin numbering system; Cys₂₉₂ in MTSP9) with a serine. Two overlapping gene fragments, each containing the AGT codon
20 for serine at position 122 were PCR amplified using the following primers: for the 5' gene fragment, TCTCTCGAGAAAAGAATAGCATCTGGAGTCATTGCACCC (SEQ ID NO. 13) and AGAGGCTTCTGGCAAATAATCCGGCGTATGTC (SEQ ID NO. 14); for the 3' gene fragment, ATTCGCGGCCGCTTAGATGCCTGTTTTTGAAGCAAT (SEQ ID NO. 21) and
25 GACATACGCCGGATTAGTTTGCCAGAAGCCTCT (SEQ ID NO. 22). The amplified gene fragments were purified on a 1% agarose gel, mixed and reamplified by PCR to produce the full length coding sequence for the protease domain of MTSP9 C122S. This sequence was then cut with restriction enzymes NotI and XhoI, and ligated into vector pPic9KX.

30 MTSP9 fermentation and initial product recovery

Fermentation

-152-

P. pastoris clone GS115/pPIC9K:MTSP9 C122S Sac MC2 expressing the C122S mutant form of MTSP9 was fermented at the 5 liter scale. An overnight culture of 200 ml (OD600 of approximately 25) was used to inoculate 3.2 liters of fermentation medium in each of four Bioflo vessels (New Brunswick Scientific, Edison, NJ). The batch phase complex medium contained 10 g/l yeast extract, 20 g/l peptone, 40 g/l glycerol, 5 g/l ammonium sulfate, 0.2 g/l calcium sulfate(dihydrate), 2 g/l magnesium sulfate(heptahydrate), 2 g/l potassium sulfate, 25 g/l sodium hexametaphosphate, and 4.35 ml/l PTM1 (6.0 g/l $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$, 0.08 g/l NaI, 3.0 g/l $\text{MnSO}_4 \cdot \text{H}_2\text{O}$, 0.2 g/l $\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$, 0.02 g/l H_3BO_3 , 0.5 g/l CoCl_2 , 20.0 g/l ZnCl_2 , 65.0 g/l $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$, 0.2 g/l biotin, 5.0 ml/l H_2SO_4). The culture was grown at a pH of 5.0 and a temperature of 28 °C in the batch phase. Concentrated ammonium hydroxide was used to maintain the pH of the culture. KFO 880 (KABO Chemicals, Cheyenne, WY) was used as needed to control foaming (see, *e.g.*, Zhang *et al.* (2000) Modeling *Pichia pastoris* Growth on Methanol and Optimizing the Production of a Recombinant Protein, the Heavy-Chain Fragment C of Botulinum Neurotoxin, Serotype A. Biotechnology and Bioengineering Vol. 70, No 1).

The batch phase of the fermentation lasted about 22 hours at which time the culture had consumed all of the initial glycerol in the medium. A substrate limited fed-batch of 50% (w/v) glycerol was initiated at 18 ml/l*hr at this point. Two hours into the glycerol fed-batch the pH of the culture was linearly increased from 5.0 to 7.0 over a two hour period by concentrated ammonium hydroxide addition. The glycerol fed-batch was about 4.5 hours in duration. The cultures reached densities of 220-250 g/l wet cell weight by this point.

Methanol induction was initiated following the end of the glycerol fed-batch phase. The culture was transitioned to methanol utilization by the method of Zhang *et al.* by adding 1.5 ml of methanol per liter of culture and linearly decreasing the glycerol feed rate from 18 ml/l*hr to 0 ml/l*hr over a 3 hour period. The methanol addition served as an on-line calibration of the MeOH Sensor (Raven Biotech, Vancouver, BC, Canada) used to control the fermenter throughout induction. After the initial amount methanol was utilized, as indicated by the MeOH Sensor, another 1.5 ml/l was added to the culture and the MeOH

-153-

sensor was used to control the methanol concentration in the fermenter at that level throughout the induction phase. The methanol fed to the fermenter was supplemented with 2 ml/l PTM4 solution (2.0 g/l $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$, 0.08 g/l NaI, 3.0 g/l $\text{MnSO}_4 \cdot \text{H}_2\text{O}$, 0.2 g/l $\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$, 0.02 g/l H_3BO_3 , 0.5 g/l $\text{CoCl}_2 \cdot 6\text{H}_2\text{O}$, 7.0 g/l ZnCl_2 , 22.0 g/l $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$, 0.2 g/l biotin, 1.0 ml/l H_2SO_4). The induction phase lasted about 42.5 hours.

Initial Product Recovery

The supernatant from each of the fermentations was harvested by centrifugation, pooled, and then was concentrated to about 0.5 liter using a 10 kDa ultrafiltration cartridge (A/G Technologies Corp., Needham, MA) on a SRT5 ultrafiltration system (North Carolina SRT, Cary, NC). The concentrate was drained from the system, then the system was rinsed with a volume of 50 mM Hepes, pH 7.0 equal to the concentrated material. The concentrate and the rinse material were combined to yield the final ultrafiltration product of about 1 liter. A final clarification of the supernatant was done with a SartoBran 300 0.45 + 0.2 μm capsule filter (Sartorius Separations Div., Edgewood, NJ).

Protein Purification – MTSP9

Concentrated fermentation supernatant of glycosylated MTSP9 was dialyzed against 50mM HEPES, pH 7.0, filtered and loaded directly onto 147mL SP Sepharose cation exchange column (Amersham-Pharmacia Biotech) which was pre-equilibrated with 50mM HEPES, pH 7.0. Protein was eluted using a linear gradient of 0-500mM NaCl over 7 column volumes at a flow rate of 5mL/min.

Active fractions were pooled then dialyzed against 50mM Na_2HPO_4 , pH 5.5 overnight. Purified, glycosylated MTSP9 was then deglycosylated by the addition 0.1 ml of Endoglycosidase H (ProZyme, 5 U/ml) per mg of protein and incubating overnight at 4°C with gentle swirling. The dialyzed protein solution was then adjusted to pH 7, filtered and and loaded directly onto 147mL SP Sepharose cation exchange column and eluted as described above. Active fractions were pooled and benzamidine was added to a final concentration of 10 mM. Protein purity was examined by SDS-PAGE and protein concentration

-154-

determined by measurement of OD₂₈₀ and use of a theoretical extinction coefficient of 2.017 ml/(mg x OD₂₈₀).

EXAMPLE 3

Assays for identification of candidate compounds that modulate that activity of an MTSP

Assay for identifying inhibitors

The ability of test compounds to act as inhibitors of catalytic activity of an MTSP9 can be assessed in an amidolytic assay. The inhibitor-induced inhibition of amidolytic activity by a recombinant MTSP or the protease domain portions thereof, can be measured by IC₅₀ values in such an assay.

The protease domain of MTSP9 expressed as described above is assayed in Costar 96 well tissue culture plates (Corning NY) for inhibition by various test compounds as follows . Approximately 1-10 nM protease is added without inhibitor, or with 100000 nM inhibitor and seven 1:6 dilutions into 1X direct buffer (29.2 mM Tris, pH 8.4, 29.2 mM Imidazole, 217 mM NaCl (100 μ L final volume)), and allowed to incubate at room temperature for 30 minutes. 400 μ M substrate S 2366 (L-pyroglutamyl-L-prolyl-L-arginine-p-nitroaniline hydrochloride; DiaPharma, Westchester, OH) is added and the reaction is monitored in a SpectraMAX Plus microplate reader (Molecular Devices, Sunnyvale CA) by following change in absorbance at 405 nm for 20 minutes at 37°C.

Identification of substrates

Particular substrates for use in the assays can be identified empirically by testing substrates. The following list of substrates are exemplary of those that can be tested.

Substrate name	Structure
S 2366	pyroGlu-Pro-Arg-pNA.HCl
spectrozyme t-PA	CH ₃ SO ₂ -D-HHT-Gly-Arg-pNA.AcOH
N-p-tosyl-Gly-Pro-Arg-pNA	N-p-tosyl-Gly-Pro-Arg-pNA
Benzoyl-Val-Gly-Arg-pNA	Benzoyl-Val-Gly-Arg-pNA
Pefachrome t-PA	CH ₃ SO ₂ -D-HHT-Gly-Arg-pNA
S 2765	N- α -Z-D-Arg-Gly-Arg-pNA.2HCl
S 2444	pyroGlu-Gly-Arg-pNA.HCl
S 2288	H-D-Ile-Pro-Arg-pNA.2HCl
spectrozyme UK	Cbo-L-(γ)Glu(α -t-BuO)-Gly-Arg-pNA.2AcOH
S 2302	H-D-Pro-Phe-Arg-pNA.2HCl

-155-

5	S 2266	H-D-Val-Leu-Arg-pNA.2HCl
	S 2222	Bz-Ile-Glu(g-OR)-Gly-Arg-pNA.HCl R = H(50%) and R = CH ₃ (50%)
	Chromozyme PK	Benzoyl-Pro-Phe-Arg-pNA
	S 2238	H-D-Phe-Pip-Arg-pNA.2HCl
	S 2251	H-D-Val-Leu-Lys-pNA.2HCl
10	Spectrozyme PI	H-D-Nle-HHT-Lys-pNA.2AcOH
		Pyr-Arg-Thr-Lys-Arg-AMC
		H-Arg-Gln-Arg-Arg-AMC
		Boc-Gln-Gly-Arg-AMC
		Z-Arg-Arg-AMC
15	Spectrozyme THE	H-D-HHT-Ala-Arg-pNA.2AcOH
	Spectrozyme fXIIa	H-D-CHT-Gly-Arg-pNA.2AcOH
		CVS 2081-6 (MeSO ₂ -dPhe-Pro-Arg-pNA)
		Pefachrome fVIIa (CH ₃ SO ₂ -D-CHA-But-Arg-pNA)

pNA = para-nitranilide (chromogenic)

AMC = amino methyl coumarin (fluorescent)

If none of the above substrates are cleaved, a coupled assay, described above, can be used. Briefly, test the ability of the protease to activate and enzyme, such as plasminogen and trypsinogen. To perform these assays, the single chain protease is incubated with a zymogen, such as plasminogen or trypsinogen, in the presence of the a known substrate, such, lys-plasminogen, for the zymogen. If the single chain activates the zymogen, the activated enzyme, such as plasmin and trypsin, will degrade the substrate therefor.

MTSP-9 Assay for screening modulators

The protease domain of MTSP9 expressed in *Pichia pastoris* was assayed for inhibition by various test compounds in Costar 96 well tissue culture plates (Corning NY). Approximately 1-20 nM MTSP9 was added without inhibitor, or with 100000 nM inhibitor and 7 1:6 dilutions to 1X direct buffer (29.2 mM Tris, pH 8.4, 29.2 mM Imidazole, 217 mM NaCl (100 μ L final volume)), and allowed to incubate at room temperature for 30 minutes. 400 μ M substrate pefachrome FVIIa (Pentapharm,, Norwalk, CT) was added and the reaction was monitored in a SpectraMAX Plus microplate reader (Molecular Devices, Sunnyvale CA) by following change in absorbance at 405 nm for 20 minutes at 37°C.

EXAMPLE 4**Other Assays**

These assays are described with reference to MTSP1, but such assays can be readily adapted for use with MTSP9.

5 **Amidolytic Assay for Determining Inhibition of Serine Protease Activity of Matriptase or MTSP1**

The ability of test compounds to act as inhibitors of rMAP catalytic activity was assessed by determining the inhibitor-induced inhibition of amidolytic activity by the MAP, as measured by IC_{50} values. The assay buffer was HBSA (10 mM Hepes, 150mM sodium chloride, pH 7.4, 0.1% bovine serum albumin). All reagents were from Sigma Chemical Co. (St. Louis, MO), unless otherwise indicated.

Two IC_{50} assays (a) one at either 30-minutes or 60-minutes (a 30-minute or a 60-minute preincubation of test compound and enzyme) and (b) one at 15 0-minutes (no preincubation of test compound and enzyme) were conducted. For the IC_{50} assay at either 30-minutes or 60-minutes, the following reagents were combined in appropriate wells of a Corning microtiter plate: 50 microliters of HBSA, 50 microliters of the test compound, diluted (covering a broad concentration range) in HBSA (or HBSA alone for uninhibited velocity measurement), and 50 microliters of the rMAP (Corvas International) diluted in 20 buffer, yielding a final enzyme concentration of 250 pM as determined by active site filtration. Following either a 30-minute or a 60-minute incubation at ambient temperature, the assay was initiated by the addition of 50 microliters of the substrate S-2765 (N- α -Benzyloxycarbonyl-D-arginyl-L-glycyl-L-arginine-p- 25 nitroaniline dihydrochloride; DiaPharma Group, Inc.; Franklin, OH) to each well, yielding a final assay volume of 200 microliters and a final substrate concentration of 100 μ M (about 4-times K_m). Before addition to the assay mixture, S-2765 was reconstituted in deionized water and diluted in HBSA. For the IC_{50} assay at 0 minutes; the same reagents were combined: 50 microliters of 30 HBSA, 50 microliters of the test compound, diluted (covering the identical concentration range) in HBSA (or HBSA alone for uninhibited velocity measurement), and 50 microliters of the substrate S-2765. The assay was

-157-

initiated by the addition of 50 microliters of rMAP. The final concentrations of all components were identical in both IC_{50} assays (at 30- or 60- and 0-minute).

The initial velocity of chromogenic substrate hydrolysis was measured in both assays by the change of absorbance at 405 nm using a Thermo Max®

- 5 Kinetic Microplate Reader (Molecular Devices) over a 5 minute period, in which less than 5% of the added substrate was used. The concentration of added inhibitor, which caused a 50% decrease in the initial rate of hydrolysis was defined as the respective IC_{50} value in each of the two assays (30- or 60-minutes and 0-minute).

10 ***In vitro* enzyme assays for specificity determination**

- The ability of compounds to act as a selective inhibitor of matriptase activity was assessed by determining the concentration of test compound that inhibits the activity of matriptase by 50%, (IC_{50}) as described in the above Example, and comparing IC_{50} value for matriptase to that determined for all or
15 some of the following serine proteases: thrombin, recombinant tissue plasminogen activator (rt-PA), plasmin, activated protein C, chymotrypsin, factor Xa and trypsin.

The buffer used for all assays was HBSA (10 mM HEPES, pH 7.5, 150 mM sodium chloride, 0.1% bovine serum albumin).

- 20 The assay for IC_{50} determinations was conducted by combining in appropriate wells of a Corning microtiter plate, 50 microliters of HBSA, 50 microliters of the test compound at a specified concentration (covering a broad concentration range) diluted in HBSA (or HBSA alone for V_0 (uninhibited velocity) measurement), and 50 microliters of the enzyme diluted in HBSA. Following a
25 30 minute incubation at ambient temperature, 50 microliters of the substrate at the concentrations specified below were added to the wells, yielding a final total volume of 200 microliters. The initial velocity of chromogenic substrate hydrolysis was measured by the change in absorbance at 405 nm using a Thermo Max® Kinetic Microplate Reader over a 5 minute period in which less
30 than 5% of the added substrate was used. The concentration of added inhibitor which caused a 50% decrease in the initial rate of hydrolysis was defined as the IC_{50} value.

-158-

Thrombin (fIIa) Assay

Enzyme activity was determined using the chromogenic substrate, Pefachrome t-PA (CH_3SO_2 -D-hexahydrotyrosine-glycyl-L-Arginine-p-nitroaniline, obtained from Pentapharm Ltd.). The substrate was reconstituted in deionized water prior to use. Purified human α -thrombin was obtained from Enzyme Research Laboratories, Inc. The buffer used for all assays was HBSA (10 mM HEPES, pH 7.5, 150 mM sodium chloride, 0.1% bovine serum albumin).

IC_{50} determinations were conducted where HBSA (50 μL), α -thrombin (50 μL) (the final enzyme concentration is 0.5 nM) and inhibitor (50 μL) (covering a broad concentration range), were combined in appropriate wells and incubated for 30 minutes at room temperature prior to the addition of substrate Pefachrome-t-PA (50 μL) (the final substrate concentration is 250 μM , about 5 times K_m). The initial velocity of Pefachrome t-PA hydrolysis was measured by the change in absorbance at 405 nm using a Thermo Max[®] Kinetic Microplate Reader over a 5 minute period in which less than 5% of the added substrate was used. The concentration of added inhibitor which caused a 50% decrease in the initial rate of hydrolysis was defined as the IC_{50} value.

Factor Xa

Factor Xa catalytic activity was determined using the chromogenic substrate S-2765 (N-benzyloxycarbonyl-D-arginine-L-glycine-L-arginine-p-nitroaniline), obtained from DiaPharma Group (Franklin, OH). All substrates were reconstituted in deionized water prior to use. The final concentration of S-2765 was 250 μM (about 5-times K_m). Purified human Factor X was obtained from Enzyme Research Laboratories, Inc. (South Bend, IN) and Factor Xa (FXa) was activated and prepared from it as described [Bock, P.E., Craig, P.A., Olson, S.T., and Singh, P. *Arch. Biochem. Biophys.* 273:375-388 (1989)]. The enzyme was diluted into HBSA prior to assay in which the final concentration was 0.25 nM.

Recombinant tissue plasminogen activator (rt-PA) Assay

rt-PA catalytic activity was determined using the substrate, Pefachrome t-PA (CH_3SO_2 -D-hexahydrotyrosine-glycyl-L-arginine-p-nitroaniline, obtained from Pentapharm Ltd.). The substrate was made up in deionized water followed by dilution in HBSA prior to the assay in which the final concentration was 500

-159-

micromolar (about 3-times K_m). Human rt-PA (Activase®) was obtained from Genentech Inc. The enzyme was reconstituted in deionized water and diluted into HBSA prior to the assay in which the final concentration was 1.0 nM.

Plasmin Assay

- 5 Plasmin catalytic activity was determined using the chromogenic substrate, S-2366 (L-pyroglutamyl-L-prolyl-L-arginine-p-nitroaniline hydrochloride), which was obtained from DiaPharma group. The substrate was made up in deionized water followed by dilution in HBSA prior to the assay in which the final concentration was 300 micromolar (about 2.5-times K_m).
- 10 Purified human plasmin was obtained from Enzyme Research Laboratories, Inc. The enzyme was diluted into HBSA prior to assay in which the final concentration was 1.0 nM.

Activated Protein C (aPC) Assay

- aPC catalytic activity was determined using the chromogenic substrate, Pefachrome PC (delta-carbobenzloxy-D-lysine-L-prolyl-L-arginine-p-nitroaniline dihydrochloride), obtained from Pentapharm Ltd.). The substrate was made up in deionized water followed by dilution in HBSA prior to the assay in which the final concentration was 400 micromolar (about 3-times K_m). Purified human aPC was obtained from Hematologic Technologies, Inc. The enzyme was diluted
- 15 into HBSA prior to assay in which the final concentration was 1.0 nM.
- 20

Chymotrypsin Assay

- Chymotrypsin catalytic activity was determined using the chromogenic substrate, S-2586 (methoxy-succinyl-L-arginine-L-prolyl-L-tyrosyl-p-nitroanilide), which was obtained from DiaPharma Group. The substrate was made up in deionized water followed by dilution in HBSA prior to the assay in which the final concentration was 100 micromolar (about 9-times K_m). Purified (3X-crystallized; CDI) bovine pancreatic alpha-chymotrypsin was obtained from Worthington Biochemical Corp. The enzyme was reconstituted in deionized water and diluted into HBSA prior to assay in which the final concentration was 0.5 nM.
- 25

-160-

Trypsin Assay

Trypsin catalytic activity was determined using the chromogenic substrate, S-2222 (benzoyl-L-isoleucine-L-glutamic acid-[gamma-methyl ester]-L-arginine-p-nitroanilide), which was obtained from DiaPharma Group. The
5 substrate was made up in deionized water followed by dilution in HBSA prior to the assay in which the final concentration was 250 micromolar (about 4-times K_m). Purified (3X-crystallized; TRL3) bovine pancreatic trypsin was obtained from Worthington Biochemical Corp. The enzyme was reconstituted in deionized water and diluted into HBSA prior to assay in which the final concentration was
10 0.5 nM.

Since modifications will be apparent to those of skill in this art, it is intended that this invention be limited only by the scope of the appended claims.

-161-

WHAT IS CLAIMED IS:

1. A substantially purified single or two chain polypeptide, comprising the protease domain of a type-II membrane-type serine protease 9 (MTSP9) or a catalytically active portion thereof.
- 5 2. The polypeptide of claim 1 that is an activated two chain protein.
3. A polypeptide of claim 1 selected from the group consisting of a polypeptide that comprises a sequence of amino acids encoded by the sequence of nucleotides set forth as nucleotides 31-729 in SEQ ID No. 5 and includes at least about 85% amino acid sequence identity with
10 the sequence of amino acids set forth in SEQ ID No. 18;
a polypeptide that comprises a sequence of amino acids encoded by the sequence of nucleotides set forth in SEQ ID No. 17;
a polypeptide that comprises a sequence of amino acids encoded by a sequence of nucleotides that hybridizes along at
15 least 70% of its full-length under conditions of high stringency to the sequence of nucleotides set forth as nucleotides 31-729 in SEQ ID No. 5 or as SEQ ID No. 17;
a polypeptide that comprises the sequence of amino acids set forth as amino acids 11-242 or SEQ ID No. 18;
20 a polypeptide that comprises a sequence of amino acids having at least about 90% sequence identity with the sequence of amino acids set forth in SEQ ID No. 6 or 18; and
a polypeptide that is encoded by a sequence of nucleotides that is a splice variant of the sequence set forth in SEQ ID No. 17.
- 25 4. The polypeptide of claim 1, wherein:
the MTSP9 portion of the polypeptide consists essentially of the protease domain of the MTSP9 or a catalytically active portion thereof.
5. The substantially purified polypeptide of claim 1, wherein the MTSP9 is a human polypeptide.
- 30 6. The substantially purified polypeptide of claim 1 that consists essentially of the protease domain of MTSP9 or a catalytically active portion of the protease domain of MTSP9.

-162-

7. The substantially purified polypeptide of claim 3 that consists essentially of the protease domain of MTSP9 or a catalytically active portion of the protease domain of MTSP9.

8. The substantially purified polypeptide of claim 1 that comprises the
5 sequence of amino acids set forth as amino acids 11-242 in SEQ ID No. 6.

9. The substantially purified polypeptide of claim 1 that comprises the sequence of amino acids set forth in SEQ ID No. 18.

10. The substantially purified polypeptide of claim 1, wherein the protease domain comprises the sequence of amino acids set forth as amino acids
10 11-242 of SEQ ID No. 6.

11. The substantially purified polypeptide of claim 1 that has more than about 80% sequence identity with a polypeptide that comprises the sequence of amino acids set forth as SEQ ID No. 6 or as the sequence of amino acids set forth as SEQ ID No. 18, wherein the polypeptide is a protease.

12. A polypeptide of claim 1, wherein the protease domain portion is
15 encoded by a nucleic acid molecule that hybridizes under conditions of high stringency along at least 70% of its full-length to a nucleic acid molecule comprising a sequence of nucleotides set forth as nucleotides 31-729 in SEQ ID No. 5 or as SEQ ID No. 17. or at least one domain thereof or a catalytically
20 active portion of the domain.

13. The polypeptide of claim 12, wherein the domain is the protease domain.

14. The polypeptide of claim 1, wherein:
the polypeptide does not comprise the complete sequence set forth in
25 SEQ ID No. 18 and includes at least amino acids 85-87 and/or 160-165 of SEQ ID No. 18.

15. A polypeptide of claim 3 that is a mutein, wherein:
up to about 50% of the amino acids are replaced with another amino acid;
30 and the resulting polypeptide is a single chain or two chain polypeptide that has catalytic activity of at least 10% of the unmutated polypeptide.

-163-

16. The polypeptide of claim 15, wherein up to about 10% of the amino acids are replaced with another amino acid.

17. The polypeptide of claim 15, wherein the resulting polypeptide is a single chain or two chain polypeptide and has catalytic activity of at least 50% of the unmutated polypeptide.

18. The polypeptide of claim 15, wherein a free Cysteine in the protease domain is replaced with another amino acid.

19. The polypeptide of claim 18, wherein the replacing amino acid is a serine.

20. An isolated substantially pure polypeptide that consists essentially of the protease domain of MTSP9.

21. A nucleic acid molecule, comprising a sequence of nucleotides that encodes the polypeptide of any of claims 1-20.

22. The nucleic acid molecule of claim 21 that comprises a sequence of nucleotides selected from the group consisting of:

(a) a sequence of nucleotides set forth as nucleotides 31-729 in SEQ ID No. 5 or as SEQ ID No. 17;

(b) a sequence of nucleotides that hybridizes under high stringency along its length or along at least about 70% of the full-length to the sequence of nucleotides set forth as nucleotides 31-729 in SEQ ID No. 5 or as SEQ ID No. 17;

(c) a sequence of nucleotides that encodes the polypeptide of SEQ ID No. 16;

(d) a sequence of nucleotides that is a splice variant of (a), (b) or (c);

(e) a sequence of nucleotides that encodes the protease domain or a catalytically active portion thereof thereof that includes a sequence of nucleotides having at least about 60%, 70%, 80%, 90% or 95% sequence identity the the sequence set forth in SEQ ID Nos. 5, 15 or 17; and and

(f) a sequence of nucleotides comprising degenerate codons of (a), (b), (c), (d) or (e).

23. An isolated nucleic molecule that encodes a mutein of claim 15.

-164-

24. A vector comprising the nucleic acid molecule of claim 21.
25. The vector of claim 24 that is an expression vector.
26. The vector of claim 24 that is a eukaryotic vector.
27. The vector of claim 25 that includes a sequence of nucleotides
5 that directs secretion of any polypeptide encoded by a sequence of nucleotides
operatively linked thereto.
28. The vector of claim 24 that is a *Pichia* vector or an *E. coli* vector.
29. A cell, comprising the vector of claim 24.
30. The cell of claim 29 that is a prokaryotic cell.
- 10 31. The cell of claim 29 that is a eukaryotic cell.
32. The cell of claim 29 that is selected from among a bacterial cell, a
yeast cell, a plant cell, an insect cell and an animal cell.
33. The cell of claim 29 that is a mammalian cell.
34. A nucleic acid molecule encoding a polypeptide of claim 6.
- 15 35. A vector, comprising nucleic acid molecule of claim 34.
36. A cell, comprising the vector of claim 35.
37. A recombinant non-human animal, wherein an endogenous gene
that encodes a polypeptide of claim 1 has been deleted or inactivated by
homologous recombination or insertional mutagenesis of the animal or an
20 ancestor thereof.
38. A method for producing a polypeptide that contains a protease
domain of an MTSP9 polypeptide, comprising:
culturing the cell of claim 29 under conditions whereby the encoded
polypeptide is expressed by the cell; and
25 recovering the expressed polypeptide.
39. The method of claim 38, wherein the polypeptide is secreted into
the culture medium.
40. The method of claim 38, wherein the cell is a *Pichia* cell.
41. The method of claim 38, wherein the polypeptide is expressed in
30 the cytoplasm of the host cell.
42. A method for producing a polypeptide that contains a protease
domain of a polypeptide, comprising:

-165-

culturing the cell of claim 36 under conditions whereby the encoded polypeptide is expressed by the cell; and
recovering the expressed polypeptide.

43. An antisense nucleic acid molecule that comprises at least 14
5 contiguous nucleotides or modified nucleotides that are complementary to a contiguous sequence of nucleotides encoding the protease domain of an MTSP9 of claim 1; or

comprises at least 16 contiguous nucleotides or modified nucleotides that are complementary to a contiguous sequence of nucleotides encoding the
10 protease domain of an MTSP9 of any of claims 1-20; or

comprises at least 30 contiguous nucleotides or modified nucleotides that are complementary to a contiguous sequence of nucleotides encoding the protease domain of an MTSP9 of any of claims 1-20, wherein the antisense molecule does include nucleotides 1162-1262 of SEQ ID No. 18.

15 44. The antisense molecule of claim 43 that includes a contiguous sequence of nucleotides that is the complement of the sequence of nucleotides set forth as nucleotides 31-729 in SEQ ID No. 5 or as SEQ ID No. 17.

45. A double-stranded RNA (dsRNA) molecule that comprises at least about 21 contiguous nucleotides or modified nucleotides from the sequence of
20 nucleotides encoding an MTSP9 of any of claims 1-20.

46. An antibody that specifically binds to the single chain form and/or two-chain form of a protease domain of the polypeptide of any of claims 1-20, or a fragment or derivative of the antibody containing a binding domain thereof, wherein the antibody is a polyclonal antibody or a monoclonal antibody.

25 47. The antibody of claim 46 that inhibits the enzymatic activity of the polypeptide.

48. An antibody that specifically binds to the single chain form and/or two-chain form of a protease domain of the polypeptide of claim 3, or a fragment or derivative of the antibody containing a binding domain thereof,
30 wherein the antibody is a polyclonal antibody or a monoclonal antibody and it inhibits the enzymatic activity of the polypeptide.

-166-

49. An antibody that specifically binds to the single chain form and/or two-chain form of a protease domain of the polypeptide of claim 6, or a fragment or derivative of the antibody containing a binding domain thereof, wherein the antibody is a polyclonal antibody or a monoclonal antibody and it
5 inhibits the enzymatic activity of the polypeptide.
50. A conjugate, comprising:
a polypeptide of any of claims 1-20, and
a targeting agent linked to the polypeptide directly or via a linker.
51. The conjugate of claim 50, wherein the targeting agent permits
10 affinity isolation or purification of the conjugate;
attachment of the conjugate to a surface;
detection of the conjugate; or
targeted delivery to a selected tissue or cell.
52. A conjugate, comprising:
15 a polypeptide of claim 3; and
a targeting agent linked to the polypeptide directly or via a linker.
53. The conjugate of claim 52, wherein the targeting agent permits
affinity isolation or purification of the conjugate;
attachment of the conjugate to a surface;
20 detection of the conjugate; or
targeted delivery to a selected tissue or cell.
54. A conjugate, comprising:
a polypeptide of claim 6; and
a targeting agent linked to the polypeptide directly or via a linker.
- 25 55. The conjugate of claim 54, wherein the targeting agent permits
affinity isolation or purification of the conjugate;
attachment of the conjugate to a surface;
detection of the conjugate; or
targeted delivery to a selected tissue or cell.

-167-

56. A combination, comprising:
an agent or treatment that inhibits the catalytic activity of the polypeptide of any of claims 1-20; and
another treatment or agent selected from anti-tumor and anti-angiogenic treatments and agents.

57. The combination of claim 56, wherein the inhibitor and the anti-tumor and/or anti-angiogenic agent are formulated in a single pharmaceutical composition or each is formulated in separate pharmaceutical compositions.

58. The combination of claim 56, wherein the inhibitor is selected from antibodies and antisense oligonucleotides and double-stranded RNA (dsRNA).

59. A solid support comprising two or more polypeptides of any of claims 1-20 linked thereto either directly or via a linker.

60. The support of claim 59, wherein the polypeptides comprise an array.

61. The support of claim 59, wherein the polypeptides comprise a plurality of different protease domains.

62. A solid support comprising two or more nucleic acid molecules of claim 21 or oligonucleotides portions thereof linked thereto either directly or via a linker, wherein the oligonucleotides contain at least 16 nucleotides.

63. The support of claim 62, wherein the nucleic acid molecules comprise an array.

64. The support of claim 62, wherein the nucleic acid molecules comprise a plurality of molecules that encode different protease domains.

65. A method for identifying compounds that modulate the protease activity of a polypeptide, comprising:

contacting a polypeptide of any of claims 1-20 with a substrate that is proteolytically cleaved by the polypeptide, and, either simultaneously, before or after, adding a test compound or plurality thereof;

measuring the amount of substrate cleaved in the presence of the test compound; and

-168-

selecting compounds that change the amount of substrate cleaved compared to a control, whereby compounds that modulate the activity of the polypeptide are identified.

66. The method of claim 65, wherein the test compounds are small
5 molecules, peptides, peptidomimetics, natural products, antibodies or fragments thereof that modulate the activity of the polypeptide.

67. The method of claim 65, wherein a plurality of the test substances are screened simultaneously.

68. The method of claim 65, wherein the polypeptide consists
10 essentially of a polypeptide encoded by a sequence of nucleotides selected from the group consisting of a sequence of nucleotides that:

(a) a sequence of nucleotides set forth as nucleotides 31-729 in SEQ ID No. 5 or as SEQ ID No. 17;

(b) a sequence of nucleotides that hybridizes under high stringency along
15 its length or along at least about 70% of the full-length to the sequence of nucleotides set forth as nucleotides 31-729 in SEQ ID No. 5 or as SEQ ID No. 17;

(c) a sequence of nucleotides that encodes the polypeptide of SEQ ID No.
16;

20 (d) a sequence of nucleotides that is a splice variant of (a), (b) or (c);

(e) a sequence of nucleotides that encodes the protease domain or a
catalytically active portion thereof thereof that includes a sequence
of nucleotides having at least about 80% or 85% sequence
identity the the sequence set forth in SEQ ID Nos. 5, 15 or 17;
25 and and

(f) a sequence of nucleotides comprising degenerate codons of (a),
(b), (c), (d) or (e).

69. The method of claim 65, wherein the polypeptide consists
essentially of a polypeptide selected from the group consisting of:

30 a polypeptide that comprises a sequence of amino acids encoded
by the sequence of nucleotides set forth as nucleotides 31-729 in SEQ ID No. 5;

-169-

a polypeptide that comprises a sequence of amino acids encoded by the sequence of nucleotides set forth in SEQ ID No. 17;

a polypeptide that comprises a sequence of amino acids encoded by a sequence of nucleotides that hybridizes under conditions of high stringency
5 to the sequence of nucleotides set forth as nucleotides 31-729 in SEQ ID No. 5 or as SEQ ID No. 17;

a polypeptide that comprises the sequence of amino acids set forth as amino acids 11-242 in SEQ ID No. 16;

a polypeptide that comprises a sequence of amino acids having at
10 least about 60% sequence identity with the sequence of amino acids set forth as amino acids 11-242 in SEQ ID No. 6 or the sequence of amino acids in SEQ ID No. 18; and

a polypeptide that is encoded by a sequence of nucleotides that is a splice variant of the sequence set forth in SEQ ID No. 18.

15 70. The method of claim 65, wherein the change in the amount of substrate cleaved is assessed by comparing the amount of substrate cleaved in the presence of the test compound with the amount of substrate cleaved in the absence of the test compound.

20 71. The method of claim 67, wherein a plurality of the polypeptides are linked to a solid support, either directly or via a linker.

72. The method of claim 71, wherein the polypeptides comprise an array.

25 73. A method of identifying a compound that specifically binds to a single-chain and/or two-chain protease domain and/or to single or two-chain polypeptide and/or to a proteolytically active portion of the single or two chain form thereof of an MTSP9 polypeptide, comprising:

contacting an MTSP9 polypeptide of any of claims 1-20 or a proteolytically active portion thereof with a test compound or plurality thereof under conditions conducive to binding thereof; and either:

30 a) identifying test compounds that specifically bind to the single chain and/or two chain form of the polypeptide or to a proteolytically active portion of the single an/or two chain form thereof, or

-170-

b) identifying test compounds that inhibit binding of a compound known to bind a single chain and/or two chain form of the polypeptide or to a proteolytically active portion of the single and/or two chain form thereof, wherein the known compound is contacted with the polypeptide before, simultaneously with or after the test compound.

74. The method of claim 73, wherein the polypeptide is linked either directly or indirectly via a linker to a solid support.

75. The method of claim 73, wherein the test compounds are small molecules, peptides, peptidomimetics, natural products, antibodies or fragments thereof.

76. The method of claim 73, wherein a plurality of the test substances are screened simultaneously.

77. The method of claim 73, wherein a plurality of the polypeptides are linked to a solid support.

78. The method of claim 73, wherein the polypeptide consists essentially of a polypeptide encoded by:

(a) a sequence of nucleotides set forth as nucleotides 31-729 in SEQ ID No. 5 or as SEQ ID No. 17;

(b) a sequence of nucleotides that hybridizes under high stringency along its length or along at least about 70% of the full-length to the sequence of nucleotides set forth as nucleotides 31-729 in SEQ ID No. 5 or as SEQ ID No. 17;

(c) a sequence of nucleotides that encodes the polypeptide of SEQ ID No. 16;

(d) a sequence of nucleotides that is a splice variant of (a), (b) or (c);

(e) a sequence of nucleotides that encodes the protease domain or a catalytically active portion thereof thereof that includes a sequence of nucleotides having at least about 80% or 85% sequence identity the the sequence set forth in SEQ ID Nos. 5, 15 or 17; and and

(f) a sequence of nucleotides comprising degenerate codons of (a), (b), (c), (d) or (e).

-171-

79. A method for identifying activators of the zymogen form of an MTSP9, comprising:

contacting a zymogen form of an MTSP9 polypeptide of any of claims 1-20 or a potentially proteolytically active portion thereof with a substrate
5 of the activated form of the polypeptide;

adding a test compound, wherein the test compound is added before, after or simultaneously with the addition of the substrate; and

detecting cleavage of the substrate, thereby identifying compounds that activate the zymogen.

80. The method of claim 79, wherein the substrate is a chromogenic substrate.

81. The method of claim 79, wherein the substrate is a L-pyroglutamyl-L-prolyl-L-arginine-p-nitroaniline hydrochloride.

82. The method of claim 79, wherein the test compound is a small
15 molecule, a nucleic acid or a polypeptide.

83. A method for treating or preventing a neoplastic disease, in a mammal, comprising administering to a mammal an effective amount of an inhibitor of a polypeptide of any of claims 1-20.

84. The method of claim 83, wherein the inhibitor is an antibody that
20 specifically binds to the polypeptide, or a fragment or derivative of the antibody containing a binding domain thereof, wherein the antibody is a polyclonal antibody or a monoclonal antibody.

85. The method of claim 83, wherein the polypeptide consists essentially of a polypeptide encoded by a sequence of nucleotides that:

(a) a sequence of nucleotides set forth as nucleotides 31-729 in SEQ ID
25 No. 5 or as SEQ ID No. 17;

(b) a sequence of nucleotides that hybridizes under high stringency along its length or along at least about 70% of the full-length to the sequence of nucleotides set forth as nucleotides 31-729 in SEQ ID No. 5 or as SEQ ID No.
30 17;

(c) a sequence of nucleotides that encodes the polypeptide of SEQ ID No. 16;

-172-

- (d) a sequence of nucleotides that is a splice variant of (a), (b) or (c);
- (e) a sequence of nucleotides that encodes the protease domain or a catalytically active portion thereof thereof that includes a sequence of nucleotides having at least about 80% or 85% sequence identity the the sequence set forth in SEQ ID Nos. 5, 15 or 17; and and
- (f) a sequence of nucleotides comprising degenerate codons of (a), (b), (c), (d) or (e).

86. The method of claim 83, wherein the polypeptide comprises a polypeptide encoded by:

- (a) a sequence of nucleotides set forth as nucleotides 31-729 in SEQ ID No. 5 or as SEQ ID No. 17;
- (b) a sequence of nucleotides that hybridizes under high stringency along its length to the sequence of nucleotides set forth as nucleotides 31-729 in SEQ ID No. 5 or as SEQ ID No. 17;
- (c) a sequence of nucleotides that encodes the polypeptide of SEQ ID No. 16;
- (d) a sequence of nucleotides that is a splice variant of (a), (b) or (c); and
- (e) degenerate codons of (a), (b), (c) or (d).

87. A method of inhibiting tumor initiation, growth or progression or treating a malignant or pre-malignant condition, comprising administering an agent that inhibits activation cleavage of the zymogen form of an MTSP9 polypeptide of any of claims 1-20 or a potentially proteolytically active portion thereof or inhibits an activity of the activated form of MTSP9 or a proteolytically active portion thereof.

88. The method of claim 87, wherein the condition is a condition of the breast, cervix, prostate, lung, ovary or colon.

89. The method of claim 87, wherein the agent is an antisense oligonucleotide, double-stranded RNA (dsRNA) or an antibody.

-173-

90. The method of claim 87, further comprising administering another treatment or agent selected from anti-tumor and anti-angiogenic treatments or agents.

91. The method of claim 87, wherein the polypeptide consists
5 essentially of a polypeptide encoded by:

(a) a sequence of nucleotides set forth as nucleotides 31-729 in SEQ ID No. 5 or as SEQ ID No. 17;

(b) a sequence of nucleotides that hybridizes under high stringency along its length or along at least about 70% of the full-length to the sequence of
10 nucleotides set forth as nucleotides 31-729 in SEQ ID No. 5 or as SEQ ID No. 17;

(c) a sequence of nucleotides that encodes the polypeptide of SEQ ID No. 16;

(d) a sequence of nucleotides that is a splice variant of (a), (b) or (c);
15 (e) a sequence of nucleotides that encodes the protease domain or a catalytically active portion thereof thereof that includes a sequence of nucleotides having at least about 80% or 85% sequence identity the the sequence set forth in SEQ ID Nos. 5, 15 or 17;
and and

20 (f) a sequence of nucleotides comprising degenerate codons of (a), (b), (c), (d) or (e).

92. The method of claim 87, wherein the polypeptide comprises a polypeptide encoded by a sequence of nucleotides that:

(a) a sequence of nucleotides set forth as nucleotides 31-729 in SEQ ID
25 No. 5 or as SEQ ID No. 17;

(b) a sequence of nucleotides that hybridizes under high stringency along its length to the sequence of nucleotides set forth as nucleotides 31-729 in SEQ ID No. 5 or as SEQ ID No. 17;

(c) a sequence of nucleotides that encodes the polypeptide of SEQ ID No.
30 16;

(d) a sequence of nucleotides that is a splice variant of (a), (b) or (c);
and

-174-

(e) degenerate codons of (a), (b), (c) or (d).

93. A method of identifying a compound that binds to the single-chain and/or two-chain form of an MTSP9 polypeptide of claim and/or to a proteolytically active portion of a single-chain and/or two-chain form of an MTSP9 polypeptide of any of claims 1-20, comprising:
- contacting a test compound with both forms;
 - determining to which form the compound binds; and
 - if it binds to a form of polypeptide, further determining whether the compound has at least one of the following properties:
- (i) inhibits activation cleavage of the single-chain zymogen form of polypeptide;
 - (ii) inhibits activity of the two-chain or single-chain form; and
 - (iii) inhibits dimerization of the polypeptide.

94. A method of detecting neoplastic disease, comprising: detecting a polypeptide that comprises a polypeptide of any of claims 1-20 in a biological sample, wherein the amount, the form and/or activity detected differs from the amount the form and/or activity of polypeptide detected from a subject who does not have neoplastic disease.

95. The method of claim 94, wherein the biological sample is selected from the group consisting of blood, urine, saliva, tears, synovial fluid, sweat, interstitial fluid, sperm, cerebrospinal fluid, ascites fluid, tumor tissue biopsy and circulating tumor cells.

96. A method of identifying a compound that binds to the single-chain and/or two-chain form of a polypeptide of any of claims 1-20, comprising:
- contacting a test compound with both forms;
 - determining to which form or forms the compound binds; and
 - if it binds to a form of polypeptide, further determining whether the compound has at least one of the following properties:
- (i) inhibits activation cleavage of the single-chain zymogen form of polypeptide;
 - (ii) inhibits activity of the two-chain or single-chain form; and
 - (iii) inhibits dimerization of the polypeptide.

-175-

97. The method of claim 96, wherein the biological sample is selected from the group consisting of blood, urine, saliva, tears, synovial fluid, sweat, interstitial fluid, cerebrospinal fluid, a sperm sample, ascites fluid, tumor tissue biopsy and circulating tumor cells.

5 98. The method of claim 96, wherein both forms consist essentially of the protease domain.

99. A method of diagnosing the presence of a pre-malignant lesion, a malignancy, or other pathologic condition in a subject, comprising:

obtaining a biological sample from the subject; and
10 exposing it to a detectable agent that binds to a two-chain and/or single-chain form of an MTSP9 polypeptide, wherein the pathological condition is characterized by the presence or absence of the two-chain or single-chain form.

100. A method of monitoring tumor progression and/or therapeutic effectiveness, comprising detecting and/or quantifying the level, the form and/or
15 activity of an MTSP9 polypeptide in a body tissue or fluid sample.

101. The method of claim 100, wherein the tumor is a tumor of the breast, cervix, prostate, lung, ovary or colon.

102. The method of claim 100, wherein the body fluid is blood, urine, sweat, saliva, cerebrospinal fluid and synovial fluid.

20 103. The polypeptide of any of claims 1-20, wherein:
the polypeptide includes at least amino acids 85-87 and/or 160-165 of SEQ ID No. 18.

104. A method for identifying compounds that modulate the protease activity of an MTSP9 polypeptide, comprising:

25 contacting a polypeptide of any of claims 1-20 or a proteolytically active portion thereof with a substrate that is proteolytically cleaved by the polypeptide, and, either simultaneously, before or after, adding a test compound or plurality thereof;

measuring the amount of substrate cleaved in the presence of the test
30 compound; and

-176-

selecting compounds that change the amount of substrate cleaved compared to a control, whereby compounds that modulate the activity of the polypeptide are identified.

105. The method of claim 104, wherein the polypeptide comprises a
5 polypeptide encoded by a sequence of nucleotides that:

(a) a sequence of nucleotides set forth as nucleotides 31-729 in SEQ ID No. 5 or as SEQ ID No. 17;

(b) a sequence of nucleotides that hybridizes under high stringency along its length or along at least about 70% of the full-length to the sequence of
10 nucleotides set forth as nucleotides 31-729 in SEQ ID No. 5 or as SEQ ID No. 17;

(c) a sequence of nucleotides that encodes the polypeptide of SEQ ID No. 16;

(d) a sequence of nucleotides that is a splice variant of (a), (b) or (c);

15 (e) a sequence of nucleotides that encodes the protease domain or a catalytically active portion thereof thereof that includes a sequence of nucleotides having at least about 80% or 85% sequence identity the the sequence set forth in SEQ ID Nos. 5, 15 or 17;
and and

20 (f) a sequence of nucleotides comprising degenerate codons of (a), (b), (c), (d) or (e).

106. The method of claim 104, wherein the polypeptide consists essentially of a polypeptide encoded by a sequence of nucleotides that:

(a) a sequence of nucleotides set forth as nucleotides 31-729 in SEQ ID
25 No. 5 or as SEQ ID No. 17;

(b) a sequence of nucleotides that hybridizes under high stringency along its length or along at least about 70% of the full-length to the sequence of nucleotides set forth as nucleotides 31-729 in SEQ ID No. 5 or as SEQ ID No. 17;

30 (c) a sequence of nucleotides that encodes the polypeptide of SEQ ID No. 16;

(d) a sequence of nucleotides that is a splice variant of (a), (b) or (c);

-177-

(e) a sequence of nucleotides that encodes the protease domain or a catalytically active portion thereof that includes a sequence of nucleotides having at least about 80% or 85% sequence identity to the sequence set forth in SEQ ID Nos. 5, 15 or 17;
5 and and

(f) a sequence of nucleotides comprising degenerate codons of (a), (b), (c), (d) or (e).

107. The polypeptide of claim 14, wherein the protease domain comprises the sequence of amino acids set forth in SEQ ID No. 16.

10 108. A transgenic non-human animal, comprising heterologous nucleic acid encoding a polypeptide of any of claims 1-20.

109. A probe or primer that comprises at least 14 contiguous nucleotides or modified nucleotides that are identical to a contiguous sequence of nucleotides encoding the protease domain of an MTSP9 of any of claims 1-
15 20; or

comprises at least 16 contiguous nucleotides or modified nucleotides that are identical to a contiguous sequence of nucleotides encoding the protease domain of an MTSP9 of any of claims 1-20; or

comprises at least 30 contiguous nucleotides or modified nucleotides that
20 are identical to a contiguous sequence of nucleotides encoding the protease domain of an MTSP9 of any of claims 1-20, wherein the antisense molecule does include nucleotides 1162-1262 of SEQ ID No. 18.

-1-

SEQUENCE LISTING

<110> Madison, Edwin
Ong, Edgar O.

<120>NUCLEIC ACID MOLECULES ENCODING A TRANSMEMBRANE SERINE PROTEASE 9, THE
ENCODED POLYPEPTIDES AND METHODS BASED THEREON

<130> 24745-1615

<140> Not Yet Assigned

<141> Herewith

<150>60/279,228

<151> 03-27-01

<150> 60/291,501

<151> 05-15-01

<160> 22

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 3147

<212> DNA

<213> Homo Sapien

<220>

<221> CDS

<222> (23)...(2589)

<223> Nucleotide sequence encoding MTSP1

<300>

<301> O'Brien, T.J. and Tanimoto, H.

<308> GenBank #AR081724

<309> 2000-08-31

<310> 5,972,616

<311> 1998-02-20

<312> 1999-10-26

<400> 1

tcaagagcgg cctcggggta cc atg ggg agc gat cgg gcc cgc aag ggc gga	52
Met Gly Ser Asp Arg Ala Arg Lys Gly Gly	
1 5 10	
ggg ggc ccg aag gac ttc ggc gcg gga ctc aag tac aac tcc cgg cac	100
Gly Gly Pro Lys Asp Phe Gly Ala Gly Leu Lys Tyr Asn Ser Arg His	
15 20 25	
gag aaa gtg aat ggc ttg gag gaa ggc gtg gag ttc ctg cca gtc aac	148
Glu Lys Val Asn Gly Leu Glu Glu Gly Val Glu Phe Leu Pro Val Asn	
30 35 40	
aac gtc aag aag gtg gaa aag cat ggc ccg ggg cgc tgg gtg gtg ctg	196
Asn Val Lys Lys Val Glu Lys His Gly Pro Gly Arg Trp Val Val Leu	
45 50 55	
gca gcc gtg ctg atc ggc ctc ctc ttg gtc ttg ctg ggg atc ggc ttc	244

-2-

Ala	Ala	Val	Leu	Ile	Gly	Leu	Leu	Leu	Val	Leu	Leu	Gly	Ile	Gly	Phe	
60						65					70					
ctg	gtg	tgg	cat	ttg	cag	tac	cgg	gac	gtg	cgt	gtc	cag	aag	gtc	ttc	292
Leu	Val	Trp	His	Leu	Gln	Tyr	Arg	Asp	Val	Arg	Val	Gln	Lys	Val	Phe	
75					80					85					90	
aat	ggc	tac	atg	agg	atc	aca	aat	gag	aat	ttt	gtg	gat	gcc	tac	gag	340
Asn	Gly	Tyr	Met	Arg	Ile	Thr	Asn	Glu	Asn	Phe	Val	Asp	Ala	Tyr	Glu	
				95					100					105		
aac	tcc	aac	tcc	act	gag	ttt	gta	agc	ctg	gcc	agc	aag	gtg	aag	gac	388
Asn	Ser	Asn	Ser	Thr	Glu	Phe	Val	Ser	Leu	Ala	Ser	Lys	Val	Lys	Asp	
			110					115					120			
gcg	ctg	aag	ctg	ctg	tac	agc	gga	gtc	cca	ttc	ctg	ggc	ccc	tac	cac	436
Ala	Leu	Lys	Leu	Leu	Tyr	Ser	Gly	Val	Pro	Phe	Leu	Gly	Pro	Tyr	His	
		125					130					135				
aag	gag	tgc	gct	gtg	acg	gcc	ttc	agc	gag	ggc	agc	gtc	atc	gcc	tac	484
Lys	Glu	Ser	Ala	Val	Thr	Ala	Phe	Ser	Glu	Gly	Ser	Val	Ile	Ala	Tyr	
	140					145						150				
tac	tgg	tct	gag	ttc	agc	atc	ccg	cag	cac	ctg	gtg	gag	gag	gcc	gag	532
Tyr	Trp	Ser	Glu	Phe	Ser	Ile	Pro	Gln	His	Leu	Val	Glu	Glu	Ala	Glu	
155					160					165					170	
cgc	gtc	atg	gcc	gag	gag	cgc	gta	gtc	atg	ctg	ccc	ccg	cgg	gcg	cgc	580
Arg	Val	Met	Ala	Glu	Glu	Arg	Val	Val	Met	Leu	Pro	Pro	Arg	Ala	Arg	
				175					180					185		
tcc	ctg	aag	tcc	ttt	gtg	gtc	acc	tca	gtg	gtg	gct	ttc	ccc	acg	gac	628
Ser	Leu	Lys	Ser	Phe	Val	Val	Thr	Ser	Val	Val	Ala	Phe	Pro	Thr	Asp	
			190					195					200			
tcc	aaa	aca	gta	cag	agg	acc	cag	gac	aac	agc	tgc	agc	ttt	ggc	ctg	676
Ser	Lys	Thr	Val	Gln	Arg	Thr	Gln	Asp	Asn	Ser	Cys	Ser	Phe	Gly	Leu	
		205					210					215				
cac	gcc	cgc	ggc	gtg	gag	ctg	atg	cgc	ttc	acc	acg	ccc	ggc	ttc	cct	724
His	Ala	Arg	Gly	Val	Glu	Leu	Met	Arg	Phe	Thr	Thr	Pro	Gly	Phe	Pro	
	220					225					230					
gac	agc	ccc	tac	ccc	gct	cat	gcc	cgc	tgc	cag	tgg	gcc	ctg	cgg	ggg	772
Asp	Ser	Pro	Tyr	Pro	Ala	His	Ala	Arg	Cys	Gln	Trp	Ala	Leu	Arg	Gly	
235					240					245					250	
gac	gcc	gac	tca	gtg	ctg	agc	ctc	acc	ttc	cgc	agc	ttt	gac	ctt	gcg	820
Asp	Ala	Asp	Ser	Val	Leu	Ser	Leu	Thr	Phe	Arg	Ser	Phe	Asp	Leu	Ala	
				255					260					265		
tcc	tgc	gac	gag	cgc	ggc	agc	gac	ctg	gtg	acg	gtg	tac	aac	acc	ctg	868
Ser	Cys	Asp	Glu	Arg	Gly	Ser	Asp	Leu	Val	Thr	Val	Tyr	Asn	Thr	Leu	
			270					275					280			
agc	ccc	atg	gag	ccc	cac	gcc	ctg	gtg	cag	ttg	tgt	ggc	acc	tac	cct	916
Ser	Pro	Met	Glu	Pro	His	Ala	Leu	Val	Gln	Leu	Cys	Gly	Thr	Tyr	Pro	
		285					290					295				

-3-

ccc tcc tac aac ctg acc ttc cac tcc tcc cag aac gtc ctg ctc atc	964
Pro Ser Tyr Asn Leu Thr Phe His Ser Ser Gln Asn Val Leu Leu Ile	
300 305 310	
aca ctg ata acc aac act gag cgg cgg cat ccc ggc ttt gag gcc acc	1012
Thr Leu Ile Thr Asn Thr Glu Arg Arg His Pro Gly Phe Glu Ala Thr	
315 320 325 330	
ttc ttc cag ctg cct agg atg agc agc tgt gga ggc cgc tta cgt aaa	1060
Phe Phe Gln Leu Pro Arg Met Ser Ser Cys Gly Gly Arg Leu Arg Lys	
335 340 345	
gcc cag ggg aca ttc aac agc ccc tac tac cca ggc cac tac cca ccc	1108
Ala Gln Gly Thr Phe Asn Ser Pro Tyr Tyr Pro Gly His Tyr Pro Pro	
350 355 360	
aac att gac tgc aca tgg aac att gag gtg ccc aac aac cag cat gtg	1156
Asn Ile Asp Cys Thr Trp Asn Ile Glu Val Pro Asn Asn Gln His Val	
365 370 375	
aag gtg agc ttc aaa ttc ttc tac ctg ctg gag ccc ggc gtg cct gcg	1204
Lys Val Ser Phe Lys Phe Phe Tyr Leu Leu Glu Pro Gly Val Pro Ala	
380 385 390	
ggc acc tgc ccc aag gac tac gtg gag atc aat ggg gag aaa tac tgc	1252
Gly Thr Cys Pro Lys Asp Tyr Val Glu Ile Asn Gly Glu Lys Tyr Cys	
395 400 405 410	
gga gag agg tcc cag ttc gtc gtc acc agc aac agc aac aag atc aca	1300
Gly Glu Arg Ser Gln Phe Val Val Thr Ser Asn Ser Asn Lys Ile Thr	
415 420 425	
gtt cgc ttc cac tca gat cag tcc tac acc gac acc ggc ttc tta gct	1348
Val Arg Phe His Ser Asp Gln Ser Tyr Thr Asp Thr Gly Phe Leu Ala	
430 435 440	
gaa tac ctc tcc tac gac tcc agt gac cca tgc ccg ggg cag ttc acg	1396
Glu Tyr Leu Ser Tyr Asp Ser Ser Asp Pro Cys Pro Gly Gln Phe Thr	
445 450 455	
tgc cgc acg ggg cgg tgt atc cgg aag gag ctg cgc tgt gat ggc tgg	1444
Cys Arg Thr Gly Arg Cys Ile Arg Lys Glu Leu Arg Cys Asp Gly Trp	
460 465 470	
gcc gac tgc acc gac cac agc gat gag ctc aac tgc agt tgc gac gcc	1492
Ala Asp Cys Thr Asp His Ser Asp Glu Leu Asn Cys Ser Cys Asp Ala	
475 480 485 490	
ggc cac cag ttc acg tgc aag aac aag ttc tgc aag ccc ctc ttc tgg	1540
Gly His Gln Phe Thr Cys Lys Asn Lys Phe Cys Lys Pro Leu Phe Trp	
495 500 505	
gtc tgc gac agt gtg aac gac tgc gga gac aac agc gac gag cag ggg	1588
Val Cys Asp Ser Val Asn Asp Cys Gly Asp Asn Ser Asp Glu Gln Gly	
510 515 520	
tgc agt tgt ccg gcc cag acc ttc agg tgt tcc aat ggg aag tgc ctc	1636
Cys Ser Cys Pro Ala Gln Thr Phe Arg Cys Ser Asn Gly Lys Cys Leu	
525 530 535	

-4-

tcg Ser 540	aaa Lys	agc Ser	cag Gln	cag Gln	tgc Cys	aat Asn 545	ggg Gly	aag Lys	gac Asp	gac Asp	tgt Cys 550	ggg Gly	gac Asp	ggg Gly	tcc Ser	1684
gac Asp 555	gag Glu	gcc Ala	tcc Ser	tgc Cys	ccc Pro 560	aag Lys	gtg Val	aac Asn	gtc Val	gtc Val	act Thr	tgt Cys	acc Thr	aaa Lys	cac His 570	1732
acc Thr	tac Tyr	cgc Arg	tgc Cys	ctc Leu 575	aat Asn	ggg Gly	ctc Leu	tgc Cys	ttg Leu 580	agc Ser	aag Lys	ggc Gly	aac Asn	cct Pro 585	gag Glu	1780
tgt Cys	gac Asp	ggg Gly	aag Lys 590	gag Glu	gac Asp	tgt Cys	agc Ser	gac Asp 595	ggc Gly	tca Ser	gat Asp	gag Glu	aag Lys 600	gac Asp	tgc Cys	1828
gac Asp	tgt Cys	ggg Gly 605	ctg Leu	cgg Arg	tca Ser	ttc Phe	acg Thr 610	aga Arg	cag Gln	gct Ala	cgt Arg	gtt Val 615	gtt Val	ggg Gly	ggc Gly	1876
acg Thr 620	gat Asp	gcg Ala	gat Asp	gag Glu	ggc Gly	gag Glu 625	tgg Trp	ccc Pro	tgg Trp	cag Gln	gta Val 630	agc Ser	ctg Leu	cat His	gct Ala	1924
ctg Leu 635	ggc Gly	cag Gln	ggc Gly	cac His	atc Ile 640	tgc Cys	ggg Gly	gct Ala	tcc Ser	ctc Leu 645	atc Ile	tct Ser	ccc Pro	aac Asn	tgg Trp 650	1972
ctg Leu	gtc Val	tct Ser	gcc Ala	gca Ala 655	cac His	tgc Cys	tac Tyr	atc Ile	gat Asp 660	gac Asp	aga Arg	gga Gly	ttc Phe	agg Arg 665	tac Tyr	2020
tca Ser	gac Asp	ccc Pro	acg Thr 670	cag Gln	tgg Trp	acg Thr	gcc Ala	ttc Phe 675	ctg Leu	ggc Gly	ttg Leu	cac His	gac Asp 680	cag Gln	agc Ser	2068
cag Gln	cgc Arg	agc Ser 685	gcc Ala	cct Pro	ggg Gly	gtg Val	cag Gln 690	gag Glu	cgc Arg	agg Arg	ctc Leu	aag Lys 695	cgc Arg	atc Ile	atc Ile	2116
tcc Ser	cac His 700	ccc Pro	ttc Phe	ttc Phe	aat Asn 705	gac Asp	ttc Phe	acc Thr	ttc Phe	gac Asp 710	tat Tyr	gac Asp	atc Ile	gcg Ala	ctg Leu	2164
ctg Leu 715	gag Glu	ctg Leu	gag Glu	aaa Lys	ccg Pro 720	gca Ala	gag Glu	tac Tyr	agc Ser	tcc Ser 725	atg Met	gtg Val	cgg Arg	ccc Pro	atc Ile 730	2212
tgc Cys	ctg Leu	ccg Pro	gac Asp	gcc Ala 735	tcc Ser	cat His	gtc Val	ttc Phe	cct Pro 740	gcc Ala	ggc Gly	aag Lys	gcc Ala	atc Ile 745	tgg Trp	2260
gtc Val	acg Thr	ggc Gly	tgg Trp 750	gga Gly	cac His	acc Thr	cag Gln	tat Tyr 755	gga Gly	ggc Gly	act Thr	ggc Gly	gcg Ala	ctg Leu	atc Ile	2308
ctg Leu	caa Gln	aag Lys	ggg Gly	gag Glu	atc Ile	cgc Arg	gtc Val	atc Ile	aac Asn	cag Gln	acc Thr	acc Thr	tgc Cys	gag Glu	aac Asn	2356

-5-

765					770					775						
ctc	ctg	ccg	cag	cag	atc	acg	ccg	cgc	atg	atg	tgc	gtg	ggc	ttc	ctc	2404
Leu	Leu	Pro	Gln	Gln	Ile	Thr	Pro	Arg	Met	Met	Cys	Val	Gly	Phe	Leu	
780					785					790						
agc	ggc	ggc	gtg	gac	tcc	tgc	cag	ggg	gat	tcc	ggg	gga	ccc	ctg	tcc	2452
Ser	Gly	Gly	Val	Asp	Ser	Cys	Gln	Gly	Asp	Ser	Gly	Gly	Pro	Leu	Ser	
795					800					805					810	
agc	gtg	gag	gcg	gat	ggg	cgg	atc	ttc	cag	gcc	ggg	gtg	gtg	agc	tgg	2500
Ser	Val	Glu	Ala	Asp	Gly	Arg	Ile	Phe	Gln	Ala	Gly	Val	Val	Ser	Trp	
815					820					825						
gga	gac	ggc	tgc	gct	cag	agg	aac	aag	cca	ggc	gtg	tac	aca	agg	ctc	2548
Gly	Asp	Gly	Cys	Ala	Gln	Arg	Asn	Lys	Pro	Gly	Val	Tyr	Thr	Arg	Leu	
830					835					840						
cct	ctg	ttt	cgg	gac	tgg	atc	aaa	gag	aac	act	ggg	gta	ta	ggggccgggg		2599
Pro	Leu	Phe	Arg	Asp	Trp	Ile	Lys	Glu	Asn	Thr	Gly	Val				
845					850					855						
ccacccaaat	gtgtacacct	gcggggccac	ccatcggtcca	ccccagtggtg	cacgcctgca											2659
ggctgggagac	tggaaccgtg	actgcaccag	cgccccccaga	acatacactg	tgaactcaat											2719
ctccagggct	ccaaatctgc	ctagaaaacc	tctcgcttcc	tcagcctcca	aagtggagct											2779
gggagggtaga	aggggaggac	actggtggtt	ctactgaccc	aactgggggc	aaagggttga											2839
agacacagcc	tccccgcga	gccccaaagt	gggcccagagc	gcgtttgtgt	atatctgcct											2899
cccctgtctg	taaggagcag	cgggaaacgga	gcttcggagc	ctcctcagtg	aaggtgggtg											2959
ggctgccgga	tctgggctgt	ggggcccttg	ggccacgctc	ttgaggaagc	ccagggtcgg											3019
aggaccctgg	aaaacagacg	ggtctgagac	tgaaattgtt	ttaccagctc	ccagggtgga											3079
cttcagtggtg	tgtatttgtg	taaatgggta	aaacaattta	tttcttttta	aaaaaaaaaa											3139
aaaaaaaaaa															3147	
<210> 2																
<211> 855																
<212> PRT																
<213> Homo Sapien																
<400> 2																
Met	Gly	Ser	Asp	Arg	Ala	Arg	Lys	Gly	Gly	Gly	Gly	Pro	Lys	Asp	Phe	
1				5				10						15		
Gly	Ala	Gly	Leu	Lys	Tyr	Asn	Ser	Arg	His	Glu	Lys	Val	Asn	Gly	Leu	
			20					25					30			
Glu	Glu	Gly	Val	Glu	Phe	Leu	Pro	Val	Asn	Asn	Val	Lys	Lys	Val	Glu	
		35					40					45				
Lys	His	Gly	Pro	Gly	Arg	Trp	Val	Val	Leu	Ala	Ala	Val	Leu	Ile	Gly	
		50				55					60					
Leu	Leu	Leu	Val	Leu	Leu	Gly	Ile	Gly	Phe	Leu	Val	Trp	His	Leu	Gln	
				70						75					80	
Tyr	Arg	Asp	Val	Arg	Val	Gln	Lys	Val	Phe	Asn	Gly	Tyr	Met	Arg	Ile	
				85					90					95		
Thr	Asn	Glu	Asn	Phe	Val	Asp	Ala	Tyr	Glu	Asn	Ser	Asn	Ser	Thr	Glu	
			100					105					110			
Phe	Val	Ser	Leu	Ala	Ser	Lys	Val	Lys	Asp	Ala	Leu	Lys	Leu	Leu	Tyr	
		115					120					125				
Ser	Gly	Val	Pro	Phe	Leu	Gly	Pro	Tyr	His	Lys	Glu	Ser	Ala	Val	Thr	
		130				135					140					
Ala	Phe	Ser	Glu	Gly	Ser	Val	Ile	Ala	Tyr	Tyr	Trp	Ser	Glu	Phe	Ser	
145					150					155					160	

-6-

Ile	Pro	Gln	His	Leu	Val	Glu	Glu	Ala	Glu	Arg	Val	Met	Ala	Glu	Glu
				165					170					175	
Arg	Val	Val	Met	Leu	Pro	Pro	Arg	Ala	Arg	Ser	Leu	Lys	Ser	Phe	Val
			180					185					190		
Val	Thr	Ser	Val	Val	Ala	Phe	Pro	Thr	Asp	Ser	Lys	Thr	Val	Gln	Arg
		195					200					205			
Thr	Gln	Asp	Asn	Ser	Cys	Ser	Phe	Gly	Leu	His	Ala	Arg	Gly	Val	Glu
	210					215					220				
Leu	Met	Arg	Phe	Thr	Thr	Pro	Gly	Phe	Pro	Asp	Ser	Pro	Tyr	Pro	Ala
225					230					235				240	
His	Ala	Arg	Cys	Gln	Trp	Ala	Leu	Arg	Gly	Asp	Ala	Asp	Ser	Val	Leu
				245					250					255	
Ser	Leu	Thr	Phe	Arg	Ser	Phe	Asp	Leu	Ala	Ser	Cys	Asp	Glu	Arg	Gly
			260				265						270		
Ser	Asp	Leu	Val	Thr	Val	Tyr	Asn	Thr	Leu	Ser	Pro	Met	Glu	Pro	His
		275					280					285			
Ala	Leu	Val	Gln	Leu	Cys	Gly	Thr	Tyr	Pro	Pro	Ser	Tyr	Asn	Leu	Thr
	290					295					300				
Phe	His	Ser	Ser	Gln	Asn	Val	Leu	Leu	Ile	Thr	Leu	Ile	Thr	Asn	Thr
305					310					315					320
Glu	Arg	Arg	His	Pro	Gly	Phe	Glu	Ala	Thr	Phe	Phe	Gln	Leu	Pro	Arg
				325					330					335	
Met	Ser	Ser	Cys	Gly	Gly	Arg	Leu	Arg	Lys	Ala	Gln	Gly	Thr	Phe	Asn
			340					345					350		
Ser	Pro	Tyr	Tyr	Pro	Gly	His	Tyr	Pro	Pro	Asn	Ile	Asp	Cys	Thr	Trp
		355					360					365			
Asn	Ile	Glu	Val	Pro	Asn	Asn	Gln	His	Val	Lys	Val	Ser	Phe	Lys	Phe
	370					375					380				
Phe	Tyr	Leu	Leu	Glu	Pro	Gly	Val	Pro	Ala	Gly	Thr	Cys	Pro	Lys	Asp
385					390					395					400
Tyr	Val	Glu	Ile	Asn	Gly	Glu	Lys	Tyr	Cys	Gly	Glu	Arg	Ser	Gln	Phe
				405					410					415	
Val	Val	Thr	Ser	Asn	Ser	Asn	Lys	Ile	Thr	Val	Arg	Phe	His	Ser	Asp
			420					425					430		
Gln	Ser	Tyr	Thr	Asp	Thr	Gly	Phe	Leu	Ala	Glu	Tyr	Leu	Ser	Tyr	Asp
		435					440					445			
Ser	Ser	Asp	Pro	Cys	Pro	Gly	Gln	Phe	Thr	Cys	Arg	Thr	Gly	Arg	Cys
		450				455					460				
Ile	Arg	Lys	Glu	Leu	Arg	Cys	Asp	Gly	Trp	Ala	Asp	Cys	Thr	Asp	His
465					470					475					480
Ser	Asp	Glu	Leu	Asn	Cys	Ser	Cys	Asp	Ala	Gly	His	Gln	Phe	Thr	Cys
				485					490					495	
Lys	Asn	Lys	Phe	Cys	Lys	Pro	Leu	Phe	Trp	Val	Cys	Asp	Ser	Val	Asn
			500					505					510		
Asp	Cys	Gly	Asp	Asn	Ser	Asp	Glu	Gln	Gly	Cys	Ser	Cys	Pro	Ala	Gln
		515					520					525			
Thr	Phe	Arg	Cys	Ser	Asn	Gly	Lys	Cys	Leu	Ser	Lys	Ser	Gln	Gln	Cys
	530					535					540				
Asn	Gly	Lys	Asp	Asp	Cys	Gly	Asp	Gly	Ser	Asp	Glu	Ala	Ser	Cys	Pro
545					550					555					560
Lys	Val	Asn	Val	Val	Thr	Cys	Thr	Lys	His	Thr	Tyr	Arg	Cys	Leu	Asn
				565					570					575	
Gly	Leu	Cys	Leu	Ser	Lys	Gly	Asn	Pro	Glu	Cys	Asp	Gly	Lys	Glu	Asp
				580				585					590		
Cys	Ser	Asp	Gly	Ser	Asp	Glu	Lys	Asp	Cys	Asp	Cys	Gly	Leu	Arg	Ser
		595					600					605			
Phe	Thr	Arg	Gln	Ala	Arg	Val	Val	Gly	Gly	Thr	Asp	Ala	Asp	Glu	Gly
	610					615					620				
Glu	Trp	Pro	Trp	Gln	Val	Ser	Leu	His	Ala	Leu	Gly	Gln	Gly	His	Ile

-7-

625	Cys	Gly	Ala	Ser	Leu	630	Ile	Ser	Pro	Asn	Trp	635	Leu	Val	Ser	Ala	Ala	640	His
					645	Asp	Arg	Gly	Phe	Arg	Tyr	650	Ser	Asp	Pro	Thr	Gln	655	Trp
				660	Gly	Leu	His	Asp	Gln	Ser	Gln		Arg	Ser	Ala	Pro	Gly		
			675	Arg	Arg	Leu	Lys	Arg	Ile	Ile	Ser		His	Pro	Phe	Phe	Asn		
			690	Asp	Tyr	Asp	Ile	Ala	Leu	Leu	Glu		Glu	Leu	Glu	Lys	Pro		
705	Ala	Glu	Tyr	Ser	Ser	Met	Val	Arg	Pro	Ile	Cys		Leu	Pro	Asp	Ala	Ser		
	#				725	Gly	Lys	Ala	Ile	Trp	Val		Thr	Gly	Trp	Gly	His		
				740	Gly	Thr	Gly	Ala	Leu	Ile	Leu		Gln	Lys	Gly	Glu	Ile		
			755	Arg	Val	Ile	Asn	Gln	Thr	Thr	Cys		Glu	Asn	Leu	Leu	Pro	Gln	Gln
			770	Thr	Pro	Arg	Met	Met	Cys	Val	Gly		Phe	Leu	Ser	Gly	Gly	Val	Asp
785	Cys	Gln	Gly	Asp	Ser	Gly	Gly	Pro	Leu	Ser	Ser		Val	Glu	Ala	Asp	Gly		
				805	Gly	Val	Val	Ser	Trp	Gly	Asp		Gly	Cys	Ala	Gln			
			820	Arg	Asn	Lys	Pro	Gly	Val	Tyr	Thr		Arg	Leu	Pro	Leu	Phe	Arg	Asp
			835	Ile	Lys	Glu	Asn	Thr	Gly	Val									
			850																

<210> 3

<211> 3147

<212> DNA

<213> Homo Sapien

<220>

<221> CDS

<222> (1865)...(2590)

<223> Nucleic acid sequence of protease domain of MTSP1

<400> 3

tcaagagcgg	cctcggggta	ccatgggggag	cgatcggggcc	cgcaagggcg	gaggggggccc	60
gaaggacttc	ggcgcggggac	tcaagtacaa	ctcccggcac	gagaaaagtga	atggccttgga	120
ggaaggcgtg	gagttcctgc	cagtcaacaa	cgtcaagaag	gtggaaaagc	atggcccggg	180
gcgctgggtg	gtgctggcag	ccgtgctgat	cggcctcttc	ttgggtcttg	tggggatcgg	240
cttcctggtg	tggcatcttc	agtaccggga	cgtgcgtgtc	cagaaggtct	tcaatggcta	300
catgaggatc	acaaatgaga	atcttctgga	tgcctacgag	aactccaact	ccactgagtt	360
tgtaagcctg	gccagcaagg	tgaaggacgc	gctgaagctg	ctgtacacgc	gagtcctcatt	420
cctggggccc	taccacaagg	agtcggctgt	gacggccttc	agcgaggcca	gcgtcatcgc	480
ctactactgg	tctgagttca	gcatcccgca	gcacctggtg	gaggaggccg	agcgcgtcat	540
ggccgaggag	cgcgtagtca	tgctgcccc	gcgggcgcgc	tccctgaagt	cctttgtggt	600
cacctcagtg	gtggctttcc	ccacggactc	caaaacagta	cagaggaccc	aggacaacag	660
ctgcagcttt	ggcctgcacg	cccgcggtgt	ggagctgatg	cgcttcacca	cgcccggcct	720
cctgacacgc	ccctaccccg	ctcatgccc	ctgccagtgg	gcccctgcggg	gggacgcgga	780
ctcagtgctg	agcctcacct	tccgcagctt	tgaccttgcg	tcctgacgacg	agcgcggcag	840
cgacctggtg	acggtgtaca	acaccctgag	ccccatggag	ccccacgccc	tggtgcagtt	900
gtgtggcacc	taccctccct	cctacaacct	gaccttcac	tcctcccaga	acgtcctgct	960
catcacactg	ataaccaaca	ctgagcggcg	gcatcccggc	tttgaggcca	ccttcttcca	1020
gctgcctagg	atgagcagct	gtggaggccg	cttacgtaaa	gcccagggga	cattcaacag	1080

-8-

cccctactac	ccaggccact	acccacccaa	cattgactgc	acatggaaca	ttgaggtgcc	1140
caacaaccag	catgtgaagg	tgagcttcaa	attcttctac	ctgctggagc	ccggcgtgcc	1200
tgccgggcacc	tgccccaagg	actacgtgga	gatcaatggg	gagaaatact	gcggagagag	1260
gtcccagttc	gtcgtcacca	gcaacagcaa	caagatcaca	gttcgcttcc	actcagatca	1320
gtcctacacc	gacaccggct	tcttagctga	atacctctcc	tacgactcca	gtgacccatg	1380
cccggggcag	ttcacgtgcc	gcacggggcg	gtgtatccgg	aaggagctgc	gctgtgatgg	1440
ctggggccgac	tgcaccgacc	acagcgatga	gctcaactgc	agttgcgacg	ccggccacca	1500
gttcacgtgc	aagaacaagt	tctgcaagcc	cctcttctgg	gtctgcgaca	gtgtgaacga	1560
ctgctggagac	aacagcgacg	agcaggggtg	cagttgtccg	gcccagacct	tcaggtgttc	1620
caatgggaag	tgcctctcga	aaagccagca	gtgcaatggg	aaggacgact	gtggggacgg	1680
gtccgacgag	gcctcctgcc	ccaaggtgaa	cgctcgtcact	tgtaccaaac	acacctaccg	1740
ctgcctcaat	gggctctgct	tgagcaaggg	caaccctgag	tgtgacggga	aggaggactg	1800
tagcgacggc	tcagatgaga	aggactgcga	ctgtgggctg	cggtcattca	cgagacaggc	1860
tcgt gtt gtt ggg ggc acg gat gcg gat gag ggc gag tgg ccc tgg cag						1909
Val Val Gly Gly Thr Asp Ala Asp Glu Gly Glu Trp Pro Trp Gln						
1		5		10	15	
gta agc ctg cat gct ctg ggc cag ggc cac atc tgc ggt gct tcc ctc						1957
Val Ser Leu His Ala Leu Gly Gln Gly His Ile Cys Gly Ala Ser Leu						
	20		25		30	
atc tct ccc aac tgg ctg gtc tct gcc gca cac tgc tac atc gat gac						2005
Ile Ser Pro Asn Trp Leu Val Ser Ala Ala His Cys Tyr Ile Asp Asp						
	35		40		45	
aga gga ttc agg tac tca gac ccc acg cag tgg acg gcc ttc ctg ggc						2053
Arg Gly Phe Arg Tyr Ser Asp Pro Thr Gln Trp Thr Ala Phe Leu Gly						
	50		55		60	
ttg cac gac cag agc cag cgc agc gcc cct ggg gtg cag gag cgc agg						2101
Leu His Asp Gln Ser Gln Arg Ser Ala Pro Gly Val Gln Glu Arg Arg						
	65		70		75	
ctc aag cgc atc atc tcc cac ccc ttc ttc aat gac ttc acc ttc gac						2149
Leu Lys Arg Ile Ile Ser His Pro Phe Phe Asn Asp Phe Thr Phe Asp						
	80		85		90	95
tat gac atc gcg ctg ctg gag ctg gag aaa ccg gca gag tac agc tcc						2197
Tyr Asp Ile Ala Leu Leu Glu Leu Glu Lys Pro Ala Glu Tyr Ser Ser						
	100		105		110	
atg gtg cgg ccc atc tgc ctg ccg gac gcc tcc cat gtc ttc cct gcc						2245
Met Val Arg Pro Ile Cys Leu Pro Asp Ala Ser His Val Phe Pro Ala						
	115		120		125	
ggc aag gcc atc tgg gtc acg ggc tgg gga cac acc cag tat gga ggc						2293
Gly Lys Ala Ile Trp Val Thr Gly Trp Gly His Thr Gln Tyr Gly Gly						
	130		135		140	
act ggc gcg ctg atc ctg caa aag ggt gag atc cgc gtc atc aac cag						2341
Thr Gly Ala Leu Ile Leu Gln Lys Gly Glu Ile Arg Val Ile Asn Gln						
	145		150		155	
acc acc tgc gag aac ctc ctg ccg cag cag atc acg ccg cgc atg atg						2389
Thr Thr Cys Glu Asn Leu Leu Pro Gln Gln Ile Thr Pro Arg Met Met						
	160		165		170	175
tgc gtg ggc ttc ctc agc ggc ggc gtg gac tcc tgc cag ggt gat tcc						2437
Cys Val Gly Phe Leu Ser Gly Gly Val Asp Ser Cys Gln Gly Asp Ser						

-9-

180										185					190												
ggg	gga	ccc	ctg	tcc	agc	gtg	gag	gcg	gat	ggg	cgg	atc	ttc	cag	gcc	2485											
Gly	Gly	Pro	Leu	Ser	Ser	Val	Glu	Ala	Asp	Gly	Arg	Ile	Phe	Gln	Ala												
			195					200					205														
ggt	gtg	gtg	agc	tgg	gga	gac	ggc	tgc	gct	cag	agg	aac	aag	cca	ggc	2533											
Gly	Val	Val	Ser	Trp	Gly	Asp	Gly	Cys	Ala	Gln	Arg	Asn	Lys	Pro	Gly												
		210					215					220															
gtg	tac	aca	agg	ctc	cct	ctg	ttt	cgg	gac	tgg	atc	aaa	gag	aac	act	2581											
Val	Tyr	Thr	Arg	Leu	Pro	Leu	Phe	Arg	Asp	Trp	Ile	Lys	Glu	Asn	Thr												
	225					230					235																
ggg	gta	tag	ggg	ccg	gggc	caccc	aaatg	tgt	acac	ctg	cg	ggg	ccacc			2630											
Gly	Val	*																									
240																											
cat	cg	tcc	ac	ccagt	gtg	c	acgc	ctg	cag	g	ctgg	agact	gg	acc	gctga	ctgc	cacc	agc	2690								
g	cccc	cagaa	cata	ca	ctgt	ga	act	ca	atc	t	ccagg	gctc	caa	atc	tgc	tag	aaa	ac	cct	2750							
ct	cg	ttc	ct	cag	ct	caa	agt	gg	agct	g	gg	aggt	tag	aa	ggg	agg	aca	ct	ggt	ggtc	2810						
tact	g	accca	act	gg	ggg	gca	aag	gtt	t	gaa	g	acac	ag	c	cccc	g	cc	ag	ctg	2870							
gg	cc	gag	gc	cg	ttt	gt	gt	t	g	c	ct	gt	ct	gt	aagg	ag	cg	ag	2930								
ctt	c	g	ag	cc	t	ct	ca	gt	ga	agg	t	gg	t	gg	g	ct	g	ag	2990								
g	cc	ac	g	ct	ct	t	g	ag	ga	g	cc	ct	g	ga	aa	ac	ag	ac	gg	gt	ct	g	ag	act	3050		
g	aa	att	gt	ttt	t	ac	ag	ct	cc	c	ag	gt	t	gg	ac	t	tc	ag	t	gt	gt	gt	gt	gt	gt	3110	
a	ac	a	att	tt	t	a	t	t	t	t	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	3147
<210> 4																											
<211> 241																											
<212> PRT																											
<213> Homo Sapien																											
<400> 4																											
Val	Val	Gly	Gly	Thr	Asp	Ala	Asp	Glu	Gly	Glu	Trp	Pro	Trp	Gln	Val												
1				5				10						15													
Ser	Leu	His	Ala	Leu	Gly	Gln	Gly	His	Ile	Cys	Gly	Ala	Ser	Leu	Ile												
			20					25					30														
Ser	Pro	Asn	Trp	Leu	Val	Ser	Ala	Ala	His	Cys	Tyr	Ile	Asp	Asp	Arg												
		35					40					45															
Gly	Phe	Arg	Tyr	Ser	Asp	Pro	Thr	Gln	Trp	Thr	Ala	Phe	Leu	Gly	Leu												
	50				55						60																
His	Asp	Gln	Ser	Gln	Arg	Ser	Ala	Pro	Gly	Val	Gln	Glu	Arg	Arg	Leu												
65				70						75					80												
Lys	Arg	Ile	Ile	Ser	His	Pro	Phe	Phe	Asn	Asp	Phe	Thr	Phe	Asp	Tyr												
			85						90					95													
Asp	Ile	Ala	Leu	Glu	Leu	Glu	Lys	Pro	Ala	Glu	Tyr	Ser	Ser	Met													
			100				105					110															
Val	Arg	Pro	Ile	Cys	Leu	Pro	Asp	Ala	Ser	His	Val	Phe	Pro	Ala	Gly												
		115					120					125															
Lys	Ala	Ile	Trp	Val	Thr	Gly	Trp	Gly	His	Thr	Gln	Tyr	Gly	Gly	Thr												
	130					135					140																
Gly	Ala	Leu	Ile	Leu	Gln	Lys	Gly	Glu	Ile	Arg	Val	Ile	Asn	Gln	Thr												
145				150						155					160												
Thr	Cys	Glu	Asn	Leu	Leu	Pro	Gln	Gln	Ile	Thr	Pro	Arg	Met	Met	Cys												
			165						170					175													
Val	Gly	Phe	Leu	Ser	Gly	Gly	Val	Asp	Ser	Cys	Gln	Gly	Asp	Ser	Gly												
			180					185					190														

-10-

Gly Pro Leu Ser Ser Val Glu Ala Asp Gly Arg Ile Phe Gln Ala Gly
 195 200 205
 Val Val Ser Trp Gly Asp Gly Cys Ala Gln Arg Asn Lys Pro Gly Val
 210 215 220
 Tyr Thr Arg Leu Pro Leu Phe Arg Asp Trp Ile Lys Glu Asn Thr Gly
 225 230 235 240
 Val

<210> 5
 <211> 777
 <212> DNA
 <213> Homo Sapien

<220>
 <221> CDS
 <222> (1)..(729)
 <223> Nucleotide sequence encoding MTSP9, including protease
 domain (31-729)

<400> 5
 aaa cga gtt gtt cca tta aac gtc aac aga ata gca tct gga gtc att 48
 Lys Arg Val Val Pro Leu Asn Val Asn Arg Ile Ala Ser Gly Val Ile
 1 5 10 15

 gca ccc aag gcg gcc tgg cct tgg caa gct tcc ctt cag tat gat aac 96
 Ala Pro Lys Ala Ala Trp Pro Trp Gln Ala Ser Leu Gln Tyr Asp Asn
 20 25 30

 atc cat cag tgt ggg gcc acc ttg att agt aac aca tgg ctt gtc act 144
 Ile His Gln Cys Gly Ala Thr Leu Ile Ser Asn Thr Trp Leu Val Thr
 35 40 45

 gca gca cac tgc ttc cag aag tat aaa aat cca cat caa tgg act gtt 192
 Ala Ala His Cys Phe Gln Lys Tyr Lys Asn Pro His Gln Trp Thr Val
 50 55 60

 agt ttt gga aca aaa atc aac cct ccc tta atg aaa aga aat gtc aga 240
 Ser Phe Gly Thr Lys Ile Asn Pro Pro Leu Met Lys Arg Asn Val Arg
 65 70 75 80

 aga ttt att atc cat gag aag tac cgc tct gca gca aga gag tac gac 288
 Arg Phe Ile Ile His Glu Lys Tyr Arg Ser Ala Ala Arg Glu Tyr Asp
 85 90 95

 att gct gtt gtg cag gtc tct tcc aga gtc acc ttt tcg gat gac ata 336
 Ile Ala Val Val Gln Val Ser Ser Arg Val Thr Phe Ser Asp Asp Ile
 100 105 110

 cgc cgg att tgt ttg cca gaa gcc tct gca tcc ttc caa cca aat ttg 384
 Arg Arg Ile Cys Leu Pro Glu Ala Ser Ala Ser Phe Gln Pro Asn Leu
 115 120 125

 act gtc cac atc aca gga ttt gga gca ctt tac tat ggt ggg gaa tcc 432
 Thr Val His Ile Thr Gly Phe Gly Ala Leu Tyr Tyr Gly Gly Glu Ser
 130 135 140

-11-

```

caa aat gat ctc cga gaa gcc aga gtg aaa atc ata agt gac gat gtc 480
Gln Asn Asp Leu Arg Glu Ala Arg Val Lys Ile Ile Ser Asp Asp Val
145 150 155 160

tgc aag caa cca cag gtg tat ggc aat gat ata aaa cct gga atg ttc 528
Cys Lys Gln Pro Gln Val Tyr Gly Asn Asp Ile Lys Pro Gly Met Phe
165 170 175

tgt gcc gga tat atg gaa gga att tat gat gcc tgc agg ggt gat tct 576
Cys Ala Gly Tyr Met Glu Gly Ile Tyr Asp Ala Cys Arg Gly Asp Ser
180 185 190

ggg gga cct tta gtc aca agg gat ctg aaa gat acg tgg tat ctc att 624
Gly Gly Pro Leu Val Thr Arg Asp Leu Lys Asp Thr Trp Tyr Leu Ile
195 200 205

gga att gta agc tgg gga gat aac tgt ggt caa aag gac aag cct gga 672
Gly Ile Val Ser Trp Gly Asp Asn Cys Gly Gln Lys Asp Lys Pro Gly
210 215 220

gtc tac aca caa gtg act tat tac cga aac tgg att gct tca aaa aca 720
Val Tyr Thr Gln Val Thr Tyr Tyr Arg Asn Trp Ile Ala Ser Lys Thr
225 230 235 240

ggc atc taa ttcacgataa aagttaaaca aagaaagctg tatgcaggtc atatatgc 777
Gly Ile

```

<210> 6

<211> 242

<212> PRT

<213> Homo sapiens

<223> MTSP9, including protease domain (11-242)

<400> 6

```

Lys Arg Val Val Pro Leu Asn Val Asn Arg Ile Ala Ser Gly Val Ile
1 5 10 15

Ala Pro Lys Ala Ala Trp Pro Trp Gln Ala Ser Leu Gln Tyr Asp Asn
20 25 30

Ile His Gln Cys Gly Ala Thr Leu Ile Ser Asn Thr Trp Leu Val Thr
35 40 45

Ala Ala His Cys Phe Gln Lys Tyr Lys Asn Pro His Gln Trp Thr Val
50 55 60

Ser Phe Gly Thr Lys Ile Asn Pro Pro Leu Met Lys Arg Asn Val Arg
65 70 75 80

Arg Phe Ile Ile His Glu Lys Tyr Arg Ser Ala Ala Arg Glu Tyr Asp
85 90 95

Ile Ala Val Val Gln Val Ser Ser Arg Val Thr Phe Ser Asp Asp Ile
100 105 110

Arg Arg Ile Cys Leu Pro Glu Ala Ser Ala Ser Phe Gln Pro Asn Leu
115 120 125

```

-12-

Thr Val His Ile Thr Gly Phe Gly Ala Leu Tyr Tyr Gly Gly Glu Ser
 130 135 140
 Gln Asn Asp Leu Arg Glu Ala Arg Val Lys Ile Ile Ser Asp Asp Val
 145 150 155 160
 Cys Lys Gln Pro Gln Val Tyr Gly Asn Asp Ile Lys Pro Gly Met Phe
 165 170 175
 Cys Ala Gly Tyr Met Glu Gly Ile Tyr Asp Ala Cys Arg Gly Asp Ser
 180 185 190
 Gly Gly Pro Leu Val Thr Arg Asp Leu Lys Asp Thr Trp Tyr Leu Ile
 195 200 205
 Gly Ile Val Ser Trp Gly Asp Asn Cys Gly Gln Lys Asp Lys Pro Gly
 210 215 220
 Val Tyr Thr Gln Val Thr Tyr Tyr Arg Asn Trp Ile Ala Ser Lys Thr
 225 230 235 240
 Gly Ile

<210> 7
 <211> 35
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> 5' end primer:

<400> 7
 ggcaagcttc ccttcagtat gataacatcc atcag 35

<210> 8
 <211> 32
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> 3' end primer

<400> 8
 aatgagatac cacgtatcctt tcagatccct tg 32

<210> 9
 <211> 32
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> PCR 5' end primer

<400> 9
 cgagttgttc cattaaacgt caacagaata gc 32

-13-

<210> 10
 <211> 33
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> PCR 3' end primer

<400> 10
 gcatacagct ttctttgttt aactttttatc gtg 33

<210> 11
 <211> 9276
 <212> DNA
 <213> Pichia pastoris

<400> 11
 agatctaaca tccaaagacg aaaggttgaa tgaaaccttt ttgccatccg acatccacag 60
 gtccattctc acacataagt gccaaacgca acaggagggg atacactagc agcagaccgt 120
 tgcaaacgca ggacctccac tcctcttctc ctcaacaccc actttttgcca tcgaaaaacc 180
 agcccagtta ttggggttga ttggagctcg ctcatcccaa ttcttctctat taggctacta 240
 acaccatgac tttatttagcc tgtctatcct ggccccctcg gcgagggttca tgtttgttta 300
 tttccgaatg caacaagctc cgcattacac ccgaacatca ctccagatga gggctttctg 360
 agtgtggggg caaatagttt catgttcccc aaatggccca aaactgacag tttaaacgct 420
 gtcttggaac ctaatatgac aaaagcgtga tctcatccaa gatgaactaa gtttggttcg 480
 ttgaaatgct aacggccagt tgggtcaaaaa gaaacttcca aaagtcgcca taccgtttgt 540
 cttgtttggg attgattgac gaatgctcaa aaataatctc attaatgctt agcgcagtct 600
 ctctatcgct tctgaacccc ggtgcacctg tgccgaaacg caaatgggga aacacccgct 660
 ttttgatga ttatgcattg tctccacatt gtatgcttcc aagattctgg tgggaatact 720
 gctgatagcc taacgtttcat gatcaaaatt taactgttct aaccctact tgacagcaat 780
 atataaacag aaggaaagctg cctgtcttta aacctttttt tttatcatca ttattagctt 840
 actttcataa ttgcgactgg ttccaattga caagcttttg attttaacga cttttaacga 900
 caacttgaga agatcaaaaa acaactaatt attcgaagga tccaaacgat gagatttcct 960
 tcaattttta ctgcagtttt attcgcagca tcctccgcgt tagctgctcc agtcaacact 1020
 acaacagaag atgaaacggc acaaattccg gctgaagctg tcatcggtta ctgagattta 1080
 gaaggggatt tcgatgttgc tgttttgcca ttttccaaca gcacaaataa cgggttattg 1140
 tttataaata ctactattgc cagcattgct gctaaagaag aaggggtatc tctcgagaaa 1200
 agagaggctg aagcttacgt agaattccct agggcgcccg cgaattaatt cgccttagac 1260
 atgactgttc ctcagttcaa gttgggcact tacgagaaga ccggtcttgc tagattctaa 1320
 tcaagaggat gtcagaatgc catttgcccg agagatgcag gcttcatttt tgatactttt 1380
 ttattttgtaa cctatatagt ataggatttt ttttgtcatt ttgtttcttc tcgtacgagc 1440
 ttgctcctga tcagcctatc tcgcagctga tgaatatctt gtggtagggg tttgggaaaa 1500
 tcattcgagt ttgatgtttt tcttggtatt tcccactcct cttcagagta cagaagatta 1560
 agtgagaagt tcgtttgtgc aagcttatcg ataagcttta atgcggtagt ttatcacagt 1620
 taaattgcta acgcagtcag gcaccgtgta tgaaatctaa caatgcgctc atcgtcatcc 1680
 tcggcacccg caccctggat gctgtaggca taggcttggg tatgccggtg ctgccggggc 1740
 tcttgccggg tatcgtccat tccgacagca tcgccagtca ctatggcggtg ctgctagcgc 1800
 tatatgcgtt gatgcaattt ctatgcgcac ccgttctcgg agcactgtcc gaccgctttg 1860
 gccgcgccc agtcctgtct gttcgctac gttgagccac tatcgactac gcgatcatgg 1920
 cgaccacacc cgtcctgtgg atctatcgaa tctaataatg agttaaaatc tctaataat 1980
 taaataagtc ccagtttctc catacgaacc ttaacagcat tgccgtgagc atctagacct 2040
 tcaacagcag ccagatccat cactgcttgg ccaatatgtt tcagtcctc aggagttagc 2100
 tcttgtgaag tgatgaactt ctggaaggtt gcagtgttaa ctccgctgta ttgacgggca 2160
 tatccgtacg ttggcaaaag gtggttggtg cccgaggagt aatctccaca actctctgga 2220
 gagtaggcac caacaaacac agatccagcg tgttgtactt gatcaacata agaagaagca 2280
 ttctcgattt gcaggatcaa gtgttcagga gcgtactgat tggacatttc caaagcctgc 2340
 tcgtaggttg caaccgatag ggtttagtag tgtgcaatac acttgcgtag aatttcaacc 2400
 cttggcaact gcacagcttg gttgtgaaca gcattctcaa ttctggcaag ctccctgtct 2460
 gtcatatcga cagccaacag aatcacctgg gaatcaatac catgttcagc ttgagacaga 2520

-14-

aggtctgagg	caacgaaatc	tggatcagcg	tatttatcag	caataactag	aacttcagaa	2580
ggcccagcag	gcatgtcaat	actacacagg	gctgatgtgt	cattttgaac	catcatcttg	2640
gcagcagtaa	cgaactgggt	tcctggacca	aatattttgt	cacacttagg	aacagtttct	2700
gttccgtaag	ccatagcagc	tactgcctgg	gcgcctcctg	ctagcacgat	acacttagca	2760
ccaaccttgt	gggcaacgta	gatgacttct	ggggtaaggg	taccatcctt	cttaggtgga	2820
gatgcaaaaa	caattttctt	gcaaccagca	actttggcag	gaacacccag	catcagggga	2880
gtggaaggca	gaattgcggt	tcaccagga	atatagaggc	caactttctc	aataggtctt	2940
gcaaaacgag	agcagactac	accagggcaa	gtctcaactt	gcaacgtctc	cgttagttga	3000
gcttcatgga	atttctgtac	gttatctata	gagagatcaa	tggctctctt	aacgttatct	3060
ggcaattgca	taagttcctc	tgggaaagga	gcttctaaca	caggtgtctt	caaagcgact	3120
acaattgggt	tggcagttag	ttctaaaagg	gctttgtcac	cattttgacg	aacattgtcg	3180
tcttcaattt	tgactaattc	cataatctgt	tccgttttct	ggataggacg	acgaagggca	3240
tcagaagggg	cttcttttagg	ggccttagaa	acgtcaattt	tgcacaattc	aatacgacct	3300
ttggcatctc	ctttccttct	tttggattct	tcttttaggt	gttccttggg	gtatcctggc	3360
aagcagtctc	catcacaatc	agtgaccttt	agggacttca	tatccaggtt	tctctccacc	3420
cattcccagg	ctatatcttc	cttggcacat	ctaactaatg	caaaataaaa	taagtcagca	3480
attttagcgt	tcaacaaaac	cttggattta	gcttctgcaa	gttcatcagc	ttcctcccta	3540
gcattgtctc	tacgatccca	ttcgtcgtca	aataaccgtt	tgggtataaga	accttctgga	3600
aaacaggaag	tgcgttccaa	caaggtggct	tccatggctc	taagaccctt	tgattggcca	3660
aagcagtctc	catcacaatc	gtgacagaaa	ccaacacctg	tttgttcaac	cacaaatttc	3720
gcacctttat	accacaaaacc	caattcgata	cccagcaact	tttgagttgc	tccagatgta	3780
gcctccggaa	tagacttttt	gtgacgacga	gattggtaga	ctccagtttg	tgtccttata	3840
gccaccaaag	tagtgaatag	ggacgagtac	accaggccca	acgagtaatt	agaagagtca	3900
tcactgacag	ggaacttttt	accatcgggg	cggtcagtag	tcaaagacgc	caacaaaatt	3960
gcataaataa	tggggatttat	gacatcttca	gaaagtctcg	attcagtagt	caattgccga	4020
gtctcagaaa	aagcataaac	accagaagca	acagtgggaag	tcacatctac	caacttttgcg	4080
agtggagaag	aaaaaggcac	agttctacta	ccgccattag	tgaactttt	caaactcgccc	4140
accagggctc	tatatagatac	agcgatacta	gcattagcgg	gcaaggatgc	aacttttatca	4200
aaatctagg	ccaaaatcac	cctagcgcct	gggatcatcc	tttggacaac	tctttctgcc	4260
atcagctcct	caaatgtgtc	ttcattgata	ccattattgt	acaacttgag	caagttgtcg	4320
tcagtcgatt	gagtgaactt	ctctgtaacg	gatgactcaa	cttgcacatt	aacttgaagc	4380
gcttttccta	ccaaactcaa	gatcaggttg	tgcagctggt	cagcagcata	gggaaacacg	4440
aagggaatat	tcatacttga	ggaattatca	aactctgcaa	cacttgcgta	tgcaggtagc	4500
tatttttatt	atcagtgagt	agtcggacag	tgagtgtagt	cttgagaaat	tctgaagccg	4560
gacctgcagg	gggggggggg	cagtcatcag	gagatcctct	acgccggacg	catcgtggcc	4620
taccaggcct	gaatcgcccc	gcgctgaggt	ctgcctcgtg	aagaagggtg	tgctgactca	4680
gctttgttgt	aggtggacca	atcatccagc	cagaaagtga	gggagccacg	gttgatgaga	4740
cggttgtcgg	gaagatgcgt	gttgggtgatt	ttgaactttt	gctttgccac	ggaacggtct	4800
caaagccggc	gtcccgtcaa	gatctgatcc	ttcaactcag	caaaaagtctg	atttattcaa	4860
attctgatta	gaaaaactca	gtcagcgtaa	tgtcttgcca	gtgttacaac	caattaacca	4920
tatcaatacc	atatttttga	tcgagcatca	aatgaaactg	caattttattc	atatcaggat	4980
agttccatag	gatggcaaga	aaaagccggt	tctgtaatga	aggagaaaac	tcaccgaggc	5040
tacaacctat	taattttccc	tcctgggtatc	ggtctgcat	tccgactcgt	ccaacatcaa	5100
tgacgactga	atccggtgag	tcgtcaaaaa	taaggttatc	aagtgagaaa	tcaccatgag	5160
caggccagcc	attacgctcg	aatggcaaaa	gcttatgcat	ttctttccag	acttggtcaa	5220
gtgattgcgc	ctgagcgaga	tcatacaaat	cactcgcatac	aaccaaaccg	ttattcattc	5280
gaatcgaatg	caaccggcgc	cgaaatacgc	gatcgctggt	aaaaggacaa	ttacaaaacag	5340
caggatattc	ttctaatacc	aggaacactg	ccagcgcatac	aacaatattt	tcacctgaat	5400
atgcatcatc	aggagtacgg	tggaaatgct	ttttcccggg	gatcgcagtg	gtgagtaacc	5460
gccagtttag	tctgaccatc	ataaaatgct	tgatggctcg	aagaggcata	aattccgtca	5520
tcagaaacaa	ctctggcgca	tcattctgtaa	catcattggc	aacgctacct	ttgccatggt	5580
gcccgcacatt	atcgcgagcc	tcgggcttcc	catacaaatcg	atagattgtc	gcacctgatt	5640
atcgcggcct	cgagcaagac	catttatacc	catataaatc	agcatccatg	ttggaattta	5700
tgtttatgta	agcagacagt	gtttcccggt	gaatatggct	cataacaccc	cttgatttac	5760
aacatcagag	attttgagac	tttattgttc	atgatgatata	atttttatct	tgtgcaatgt	5820
ccggcgccac	aggtgcggtt	acaacgtggc	tttccccccc	ccccctgcag	gtcggcatca	5880
gggctcgcca	cttcgggctc	gctggcgctt	atatcgccga	catcacccgat	ggggaagatc	5940
tggccggggg	actgttgggc	atgagcgctt	gtttcggcgt	gggtatggtg	gcaggccccc	6000
		gccatctcct	tgcatgcacc	attccttgcg	gcggcggtgc	6060

-15-

tcaacggcct	caacctacta	ctgggctgct	tccaatgca	ggagtgcgat	aagggagagc	6120
gtcagtatc	tatgattgga	agtatgggaa	tggatgatacc	cgcattcttc	agtgtcttga	6180
ggtctcctat	cagattatgc	ccaactaaag	caaccggagg	aggagatttc	atggtaaatt	6240
tctctgactt	ttgggtcatca	gtagactcga	actgtgagac	tatctcgggt	atgacagcag	6300
aaatgtcctt	cttgggagaca	gtaaatgaag	tcccaccaat	aaagaaatcc	ttgttatcag	6360
gaacaaactt	cttgtttcga	actttttcgg	tgccttgaac	tataaaatgt	agagtggata	6420
tgtcgggtag	gaatggagcg	ggcaaatgct	taccttctgg	accttcaaga	ggtatgtagg	6480
gtttgtagat	actgatgcca	acttcagtga	caacggttgc	atttcgttca	aaccattccg	6540
aatccagaga	aatcaaagtt	gtttgtctac	tattgatcca	agccagtgcg	gtcttgaaac	6600
tgacaatagt	gtgctcgtgt	tttgagggtca	tctttgtatg	aataaatcta	gtctttgatc	6660
taaataatct	tgacgagcca	aggcgataaa	tacccaaatc	taaaactctt	ttaaaactgt	6720
aaaaggacaa	gtatgtctgc	ctgtattaaa	ccccaaatca	gctcgtagtc	tgatcctcat	6780
caacttgagg	ggcactatct	tgttttagag	aaatttgctg	agatgcgata	tcgagaaaaa	6840
ggtacgctga	ttttaaacgt	gaaatttatc	tcaagatctc	tgcctcgcgc	gtttcgggtga	6900
tgacggtgaa	aacctctgac	acatgcagct	cccggagacg	gtcacagctt	gtctgtaagc	6960
ggatgccggg	agcagacaag	cccgtcaggg	cgcgtcagcg	ggtgttggcg	ggtgtcgggg	7020
cgcagccatg	accagtcac	gtagcgatag	cggagtgtat	actggcttaa	ctatgcggca	7080
tcagagcaga	ttgtactgag	agtgcaccat	atgcgggtgtg	aaataccgca	cagatgcgta	7140
aggagaaaaat	accgcatcag	gcgctcttcc	gcttctctgc	tcaactgactc	gctgcgctcg	7200
gtcgttcggc	tgcggcgagc	ggtatcagct	cactcaaaag	cggtaatacg	gttatccaca	7260
gaatcagggg	ataacgcagg	aaagaacatg	tgagcaaaag	gccagcaaaa	ggccaggaac	7320
cgtaaaaagg	ccgcgttgct	ggcggttttt	cataggctcc	gccccctga	cgagcatcac	7380
aaaaatcgac	gctcaagtca	gaggtggcga	aaccgcagac	gactataaaag	ataccaggcg	7440
tttccccctg	gaagctccct	cgtgcgctct	cctgttccga	ccctgccgct	taccggatac	7500
ctgtccgcct	ttctcccttc	gggaagcgtg	gcgctttctc	aatgctcacg	ctgtaggtat	7560
ctcagttcgg	tgtaggtcgt	tcgctccaag	ctgggctgtg	tgacacgaacc	cccgttcag	7620
cccgcgctct	gcgccttatc	cggtaactat	cgtcttgagt	ccaaccgggt	aagacacgac	7680
ttatcgccac	tggcagcagc	cactggtaac	aggattagca	gagcgaggta	tgtaggcggt	7740
gctacagagt	tcttgaagtg	gtggcctaac	tacggctaca	ctagaaggac	agtatttgggt	7800
atctgcgctc	tgctgaagcc	agttaccttc	ggaaaaagag	ttggtagctc	ttgatccggc	7860
aaacaaacca	ccgctggtag	cgggtggttt	tttgtttgca	agcagcagat	tacgcgcaga	7920
aaaaaaggat	ctcaagaaga	tcttttgatc	ttttctacgg	ggtctgacgc	tcagtggaaac	7980
gaaaaactcac	gttaagggat	tttgggtcatg	agattatcaa	aaaggatctt	cacctagatc	8040
ctttttaaatt	aaaaatgaag	ttttaaatca	atctaagta	tatatgagta	aacttgggtct	8100
gacagttacc	aatgcttaat	cagtgaggca	ccatctcag	cgatctgtct	atttcgttca	8160
tccatagttg	cctgactccc	cgtcgtgtag	ataactacga	tacgggaggg	cttaccatct	8220
ggccccagtg	ctgcaatgat	accgcgagac	ccacgctcac	cggctccaga	tttatcagca	8280
ataaaccagc	cagccggaag	ggccgagcgc	agaagtgggt	ctgcaacttt	atccgcctcc	8340
atccagtcta	ttaattgttg	ccgggaagct	agagtaagta	gttcgccagt	taatagtttg	8400
cgcacgcttg	ttgccattgc	tgcaggcatc	gtggtgtcac	gctcgtcgtt	tggtagggt	8460
tcattcagct	ccggttccca	acgatcaagg	cgagttacat	gatcccccat	gttgtgcaaa	8520
aaagcggtta	gctccttcgg	tctccgatac	gttgtcagaa	gtaagtggc	cgcagtgtta	8580
tcactcatgg	ttatggcagc	actgcataat	tctcttactg	tcatgccatc	cgtaagatgc	8640
ttttctgtga	ctggtgagta	ctcaaccaag	tcattctgag	aatagtgtat	gccgacgacc	8700
agttgtctct	gcccggcgtc	aacacgggat	aataccgcgc	cacatagcag	aactttaaaa	8760
gtgctcatca	ttggaaaacg	ttcttcgggg	cgaaaactct	caaggatctt	accgctgttg	8820
agatccagtt	cgatgtaacc	cactcgtgca	cccaactgat	cttcagcatc	ttttactttc	8880
accagcgttt	ctgggtgagc	aaaaacagga	aggcaaaaatg	ccgcaaaaaa	gggaataaag	8940
gcgacacgga	aatgattgaat	actcatactc	ttcctttttc	aatattattg	aagcatttat	9000
cagggttatt	gtctcatgag	cggatacata	tttgaatgta	tttagaaaaa	taacaaaata	9060
ggggttccgc	gcacatttcc	ccgaaaagtg	ccacctgacg	tctaagaaac	cattattatc	9120
atgacattaa	cctataaaaa	taggcgtatc	acgaggccct	ttcgtcttca	agaattaatt	9180
ctcatgtttg	acagcttatc	atcgataagc	tgactcatgt	tggatttgtg	aaatagacgc	9240
agatcgggaa	cactgaaaaa	taacagttat	tattcg			9276

<210> 12
 <211> 11
 <212> PRT

-16-

<213> MTSP9 region adjacent to and including the cleavage site starting the protease domain (Arg↓Ile)

<400> 12
 Arg Val Val Pro Leu Asn Val Asn Arg Ile Ala
 1 5 10

<210> 13
 <211> 39
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Primer

<400> 13
 tctctcgaga aaagaatagc atctggagtc attgcaccc 39

<210> 14
 <211> 33
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> primer

<400> 14
 agaggcttct ggcaaactaa tcggcgtatg tc 32

<210> 15
 <211> 11
 <212> PRT
 <213> Pichia protease cleavage site
 <400> 15

Lys Arg Ile Ala Ser Gly Val Ile Ala Pro Lys
 1 5 10

<210> 16
 <211> 232
 <212> PRT
 <213> Homo sapiens
 <223> Protease domain of MTSP9 (1-232)

<400> 16
 Ile Ala Ser Gly Val Ile Ala Pro Lys Ala Ala Trp Pro Trp Gln Ala
 1 5 10 15
 Ser Leu Gln Tyr Asp Asn Ile His Gln Cys Gly Ala Thr Leu Ile Ser
 20 25 30
 Asn Thr Trp Leu Val Thr Ala Ala His Cys Phe Gln Lys Tyr Lys Asn
 35 40 45
 Pro His Gln Trp Thr Val Ser Phe Gly Thr Lys Ile Asn Pro Pro Leu

-17-

50	55	60
Met Lys Arg Asn Val Arg Arg Phe Ile Ile His Glu Lys Tyr Arg Ser		
65	70	75 80
Ala Ala Arg Glu Tyr Asp Ile Ala Val Val Gln Val Ser Ser Arg Val		
	85	90 95
Thr Phe Ser Asp Asp Ile Arg Arg Ile Cys Leu Pro Glu Ala Ser Ala		
	100	105 110
Ser Phe Gln Pro Asn Leu Thr Val His Ile Thr Gly Phe Gly Ala Leu		
	115	120 125
Tyr Tyr Gly Gly Glu Ser Gln Asn Asp Leu Arg Glu Ala Arg Val Lys		
	130	135 140
Ile Ile Ser Asp Asp Val Cys Lys Gln Pro Gln Val Tyr Gly Asn Asp		
	145	150 155 160
Ile Lys Pro Gly Met Phe Cys Ala Gly Tyr Met Glu Gly Ile Tyr Asp		
	165	170 175
Ala Cys Arg Gly Asp Ser Gly Gly Pro Leu Val Thr Arg Asp Leu Lys		
	180	185 190
Asp Thr Trp Tyr Leu Ile Gly Ile Val Ser Trp Gly Asp Asn Cys Gly		
	195	200 205
Gln Lys Asp Lys Pro Gly Val Tyr Thr Gln Val Thr Tyr Tyr Arg Asn		
	210	215 220
Trp Ile Ala Ser Lys Thr Gly Ile		
	225	230

<210> 17

<211> 1257

<212> DNA

<213> Homo sapien

<220>

<221> CDS

<222> (1)...(1257)

<223> Nucleotide sequence encoding full-length MTSP9

<400> 17

atg atg tat cgg aca gta gga ttt ggc acc cga agc aga aat ctg aag	48
Met Met Tyr Arg Thr Val Gly Phe Gly Thr Arg Ser Arg Asn Leu Lys	
1 5 10 15	

cca tgg atg att gcc gtt ctc att gtg ttg tcc ctg aca gtg gtg gca	96
Pro Trp Met Ile Ala Val Leu Ile Val Leu Ser Leu Thr Val Val Ala	
20 25 30	

gtg acc ata ggt ctc ctg gtt cac ttc cta gta ttt gac caa aaa aag	144
Val Thr Ile Gly Leu Leu Val His Phe Leu Val Phe Asp Gln Lys Lys	
35 40 45	

-18-

gag tac tat cat ggc tcc ttt aaa att tta gat cca caa atc aat aac	192
Glu Tyr Tyr His Gly Ser Phe Lys Ile Leu Asp Pro Gln Ile Asn Asn	
50 55 60	
aat ttc gga caa agc aac aca tat caa ctt aag gac tta cga gag acg	240
Asn Phe Gly Gln Ser Asn Thr Tyr Gln Leu Lys Asp Leu Arg Glu Thr	
65 70 75 80	
acc gaa aat ttg gtg gat gag ata ttt ata gat tca gcc tgg aag aaa	288
Thr Glu Asn Leu Val Asp Glu Ile Phe Ile Asp Ser Ala Trp Lys Lys	
85 90 95	
aat tat atc aag aac caa gta gtc aga ctg act cca gag gaa gat ggt	336
Asn Tyr Ile Lys Asn Gln Val Val Arg Leu Thr Pro Glu Glu Asp Gly	
100 105 110	
gtg aaa gta gat gtc att atg gtg ttc cag ttc ccc tct act gaa caa	384
Val Lys Val Asp Val Ile Met Val Phe Gln Phe Pro Ser Thr Glu Gln	
115 120 125	
agg gca gta aga gag aag aaa atc caa agc atc tta aat cag aag ata	432
Arg Ala Val Arg Glu Lys Lys Ile Gln Ser Ile Leu Asn Gln Lys Ile	
130 135 140	
agg aat tta aga gcc ttg cca ata aat gcc tca tca gtt caa gtt aat	480
Arg Asn Leu Arg Ala Leu Pro Ile Asn Ala Ser Ser Val Gln Val Asn	
145 150 155 160	
gca atg agc tca tca aca ggg gag tta act gtc caa gca agt tgt ggt	528
Ala Met Ser Ser Thr Gly Glu Leu Thr Val Gln Ala Ser Cys Gly	
165 170 175	
aaa cga gtt gtt cca tta aac gtc aac aga ata gca tct gga gtc att	576
Lys Arg Val Val Pro Leu Asn Val Asn Arg Ile Ala Ser Gly Val Ile	
180 185 190	
gca ccc aag gcg gcc tgg cct tgg caa gct tcc ctt cag tat gat aac	624
Ala Pro Lys Ala Ala Trp Pro Trp Gln Ala Ser Leu Gln Tyr Asp Asn	
195 200 205	
atc cat cag tgt ggg gcc acc ttg att agt aac aca tgg ctt gtc act	672
Ile His Gln Cys Gly Ala Thr Leu Ile Ser Asn Thr Trp Leu Val Thr	
210 215 220	
gca gca cac tgc ttc cag aag tat aaa aat cca cat caa tgg act gtt	720
Ala Ala His Cys Phe Gln Lys Tyr Lys Asn Pro His Gln Trp Thr Val	
225 230 235 240	
agt ttt gga aca aaa atc aac cct ccc tta atg aaa aga aat gtc aga	768
Ser Phe Gly Thr Lys Ile Asn Pro Pro Leu Met Lys Arg Asn Val Arg	
245 250 255	
aga ttt att atc cat gag aag tac cgc tct gca gca aga gag tac gac	816
Arg Phe Ile Ile His Glu Lys Tyr Arg Ser Ala Ala Arg Glu Tyr Asp	
260 265 270	
att gct gtt gtg cag gtc tct tcc aga gtc acc ttt tcg gat gac ata	864
Ile Ala Val Val Gln Val Ser Ser Arg Val Thr Phe Ser Asp Asp Ile	
275 280 285	

-19-

cgc cgg att tgt ttg cca gaa gcc tct gca tcc ttc caa cca aat ttg	912
Arg Arg Ile Cys Leu Pro Glu Ala Ser Ala Ser Phe Gln Pro Asn Leu	
290 295 300	
act gtc cac atc aca gga ttt gga gca ctt tac tat ggt ggg gaa tcc	960
Thr Val His Ile Thr Gly Phe Gly Ala Leu Tyr Tyr Gly Gly Glu Ser	
305 310 315 320	
caa aat gat ctc cga gaa gcc aga gtg aaa atc ata agt gac gat gtc	1008
Gln Asn Asp Leu Arg Glu Ala Arg Val Lys Ile Ile Ser Asp Asp Val	
325 330 335	
tgc aag caa cca cag gtg tat ggc aat gat ata aaa cct gga atg ttc	1056
Cys Lys Gln Pro Gln Val Tyr Gly Asn Asp Ile Lys Pro Gly Met Phe	
340 345 350	
tgt gcc gga tat atg gaa gga att tat gat gcc tgc agg ggt gat tct	1104
Cys Ala Gly Tyr Met Glu Gly Ile Tyr Asp Ala Cys Arg Gly Asp Ser	
355 360 365	
ggg gga cct tta gtc aca agg gat ctg aaa gat acg tgg tat ctc att	1152
Gly Gly Pro Leu Val Thr Arg Asp Leu Lys Asp Thr Trp Tyr Leu Ile	
370 375 380	
gga att gta agc tgg gga gat aac tgt ggt caa aag gac aag cct gga	1200
Gly Ile Val Ser Trp Gly Asp Asn Cys Gly Gln Lys Asp Lys Pro Gly	
385 390 395 400	
gtc tac aca caa gtg act tat tac cga aac tgg att gct tca aaa aca	1248
Val Tyr Thr Gln Val Thr Tyr Tyr Arg Asn Trp Ile Ala Ser Lys Thr	
405 410 415	
ggc atc taa	1257
Gly Ile *	

<210> 18
 <211> 418
 <212> PRT
 <213> Homo sapien

<400> 18
 Met Met Tyr Arg Thr Val Gly Phe Gly Thr Arg Ser Arg Asn Leu Lys
 1 5 10 15
 Pro Trp Met Ile Ala Val Leu Ile Val Leu Ser Leu Thr Val Val Ala
 20 25 30
 Val Thr Ile Gly Leu Leu Val His Phe Leu Val Phe Asp Gln Lys Lys
 35 40 45
 Glu Tyr Tyr His Gly Ser Phe Lys Ile Leu Asp Pro Gln Ile Asn Asn
 50 55 60
 Asn Phe Gly Gln Ser Asn Thr Tyr Gln Leu Lys Asp Leu Arg Glu Thr
 65 70 75 80
 Thr Glu Asn Leu Val Asp Glu Ile Phe Ile Asp Ser Ala Trp Lys Lys
 85 90 95
 Asn Tyr Ile Lys Asn Gln Val Val Arg Leu Thr Pro Glu Glu Asp Gly
 100 105 110
 Val Lys Val Asp Val Ile Met Val Phe Gln Phe Pro Ser Thr Glu Gln

-20-

[illegible]

```
<210> 19
<211> 32
<212> DNA
<213> Artificial Sequence
```

<220>
<223> Primer

```
<400> 19
aatgagatac cacgtatctt tcagatccct tg
```

32

```
<210> 20
<211> 29
<212> DNA
<213> Artificial Sequence
```

<220>
<223> Primer

-21-

<400> 20
atgagaagta ccgctctgca gcaagagag 29

<210> 21
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 21
attcgcggcc gcttagatgc ctgtttttga agcaat 36

<210> 22
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 22
gacatacgcc ggattagttt gccagaagcc tct 33